09/326402 C9C

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450 Alexandria, VA 22313 on January 14, 2005.

Frank C. Eisenschenk

Frank C. Eisenschenk, Ph.D., Patent Attorney

REQUEST FOR CERTIFICATE OF CORRECTION UNDER 37 CFR 1.322 and 1.323
Docket No. GEN-T112XC1
Patent No. 6,759,192

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov
Issued	:	July 6, 2004
Patent No.	:	6,759,192
For	:	Polymorphic Markers of the Prostate Carcinoma Tumor Antigen-1 (PCTA-1)

ATTN: CERTIFICATE OF CORRECTIONS BRANCH
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR CERTIFICATE OF CORRECTION UNDER 37 CFR 1.322 (OFFICE MISTAKE) AND UNDER 37 CFR 1.323 (APPLICANTS' MISTAKE)

Sir:

A Certificate of Correction for the above-identified patent has been prepared and is attached hereto.

In the left-hand column below is the column and line number where errors occurred in the patent. In the right-hand column is the page and line number in the application where the correct information appears.

Patent Reads:	**Application Reads:**
Column 4, Line 24: "and (consenus"	See Amendment dated 12/15/03, page 2: --and 22 (consensus)--
Column 4, Line 59: "82393. SEQ ID NO: 12"	See Examiner's Amendment dated 02/24/04, page 3: --82393.--
Column 5, Line 60: "Polynucleotides"	Page 7, Line 2: --polynucleotides--
Column 7, Line 17: "Volypeptide"	Page 8, Line 26: --polypeptide--
Column 8, Line 27: "heterozyposity"	Page 10, Line 6: --heterozygosity rate--
Column 11, Line 14: "bronchial En asthma"	Page 13, Lines 20-21: --bronchial asthma--
Column 14, Line 1: "BLASr"	Page 17, Line 8: --BLAST--
Column 15, Line 15: "exon 7, exon 9"	Page 18, Lines 29-30: --exon 7, exon 8, exon 9--
Column 20, Line 7: "nucleotide c at"	Page 24, Lines 30-31: --nucleotide C at--
Column 25, Line 8: "4400-8000"	Page 30, Line 24: --44001-48000--
Column 25, Line 44: "3'-regulatory,polynucleotide"	Page 31, Line 12: --3'-regulatory polynucleotide --

Column 27, Line 18: "seven let"	Page 33, Line 17: --seven *tet*--
Column 27, Line 34: "the let operator"	Page 33, Line 27: --the *tet* operator--
Column 31, Line 12: "24042406"	Page 38, Line 13: --2404-2406--
Column 31, Line 49: "nucleotide Tat positions"	Page 38, Line 35: --nucleotide T at positions--
Column 32, line 62: "thereto;"	Page 40, Lines 16-17: --thereto; and--
Column 33, Line 55: "A118, A123"	Page 41, Lines 20-21: --A118, and A123--
Column 34, Line 42: "A92A94to"	Page 42, Line 20: --A92, A94 to --
Column 35, Line 3: "A1107"	Page 43, Line 3: --A107--
Column 35, Line 9: "and complements"	Page 43, Lines 6-7: --and the complements--
Column 35, Line 13: "A55A56A59A92A94to"	Page 43, Line 9: --A55, A56, A59, A92, A94 to--
Column 35, Line 13: "A108A111 to A113A115"	Page 43, Line 9: --A108, A111 to A113, A115--
Column 36, Line 13: "^{32}p"	Page 44, Line 19: --^{32}P--
Column 38, Line 45: "VLSIPS*v*"	Page 47, Line 25: --VLSIPS™--

Column 50, Line 16:
"A123 complements"

Column 50, Line 20:
"A114, A122"

Column 50, Line 25:
"to 113, A115, A117, and"

Column 50, Line 64:
"A123 complements"

Column 51, Line 1:
"A109, A122,"

Column 52, Line 10:
"are Q selected"

Column 56, Line 48:
"the w biallelic"

Column 58, Line 25:
"A109, A122"

Column 58, Line 29:
"A94 to A01,"

Column 62, Line 9:
"Eli"

Column 66, Line 25:
"about IS"

Column 67, Line 27:
"WO 98120165"

Column 73, Lines 22-23:
"with overlapping"

Page 62, Lines 10-11:
--A123 to A125, and the complements"--

Page 62, Lines 13-14:
--A114, and A122--

Page 62, Lines 16-17:
--to A113, A115 to A117, and--

Page 63, Lines 6-7:
--A123 to A125, and the complements--

Page 63, Lines 9-10:
--A109, A110, A114, and A122--

Page 64, Lines 22-23:
--are selected--

Page 70, Lines 15-16:
--the biallelic--

Page 72, Lines 18-19:
--A109, A110, A114, and A122--

Page 72, Line 21:
--A94 to A101,--

Page 77, Line 14:
--E11--

Page 82, Line 29:
--about 15--

Page 84, Line 4:
--WO 98/20165--

Page 91, Lines 21-22:
--with non-overlapping--

Column 87, Line 2: "(a_i,a_j, a_i,b_j: b_i,a_j)"	Page 108, Line 19: --(a_i,a_j; a_i,b_j; b_i,a_j)--
Column 87, Line 7: "θ4=——=frequency"	Page 108, Line 22: --θ4= - - = frequency--
Column 87, Line 9: "θ3=—+=frequency"	Page 108, Line 23: --θ3= - + = frequency--
Column 87, Line 11: "θ2=+—=frequency"	Page 108, Line 24: --θ2= + - = frequency--
Column 87, Line 14: "(M_i; M_j)"	Page 108, Line 26: --(M_i, M_j)--
Column 87, Line 21: "D_{aiaj} ($2n_1$+"	Page 108, Line 30: --D_{aiaj}= ($2n_1$+--
Column 92, Line 29: "AS5"	Page 115, Line 18: --A55--
Column 95, Lines 41-42: "preferred 1 5 embodiment"	Page 119, Lines 14-15: --preferred embodiment--
Column 101, Line 49: "be I,X;1 expressed"	Page 127, Lines 13-14: --be expressed--
Column 109, Line 22: "one-genes"	Page 137, Line 10: --*onc*-genes--
Column 117, Line 15: "(ISPR)"	Page 147, Line 13: --(SPR)--
Column 117, Line 17: "carboxymethyi"	Page 147, Line 14: --carboxymethyl--
Column 118, Line 16: "(MaTa gal4"	Page 148, Line 25: --(MATa gal4--

Column 118, Line 18: "lacZmet'),""	Page 148, Line 26: --lacZmet$^-$)--
Column 118, Line 52: "GALA"	Line 149, Line 13: --GAL4--
Column 125, Line 6: "more CF particularly"	Page 157, Lines 22-23: --more particularly--
Column 127, Line 57: "monoleatc"	Page 161, Line 8: --monoleate--
Column 132, Line 42: "CAROM"	Page 167, Line 1: --CD-ROM--
Column 134, Line 47: "or a By polypeptide"	Page 169, Lines 20-21: --or a polypeptide--
Column 136, Line 47: "TTATAA Box"	See Amendment dated 02/24/03, page 2: --TAATAA Box--
Column 136, Line 48: "TATAA"	See Amendment dated 02/24/03, page 2: --TAATAA--
Column 140, Line 10: "[BLANK]"	Page 176, Line 6: --PCR assays were performed using the following protocol:--
Column 158, Line 1 of Table 6: "analysis of the"	Page 195, Line 1: --analysis for the--

Column 159, Line 2:
"analysis of the"

Page 195, Line 1:
--analysis for the--

Column 159, Line 6 of Table 6:
"

A2	A30	A41	A55	A57	A75

"

Page 195, Lines 3-5:
--

	A2	A30	A41	A55	A57	A75
frequency %	60/67 (A)	64/66 (T)	73/71 (T)	64/68 (C)	65/69 (G)	94/95 (G)
abs diff freq. all.	-7,4	-2,0	2,7	-3,7	-3,9	-1
pvalue	7,7e-03	4,3e-01	2,9e-01	1,6e-01	1,4e-01	3,4e-01
Cases/controls ↓						

--

Column 164, Line 14:
"and b dose"

Page 199, Lines 20-21:
--and dose--

Column 164, Line 34:
"they arc"

Page 199, Line 35:
--they are--

Column 167, Line 25:
"Baltimore, 1991 Ouchterlony, O."

Page 204, Lines 30-31:
--Baltimore, 1991
Ouchterlony, O.--

Column 168, Line 22:
"119-1123"

Page 206, Line 6:
--1119-1123--

Column 168, Line 26:
"Stembert"

Page 206, Line 10:
--Sternberg--

Column 168, Line 26:
"5:397404"

Page 206, Line 10:
--5:397-404--

Column 168, Line 27:
"Strinmatter"

Page 206, Line 11:
--Strittmatter--

Patent Reads:

Application Should Read:

Column 569, Line 28:
"and b) performed"

See Amendment dated 12/15/03, page 6, line 6, Claim 137:
--and b) are performed--

Column 570, Line 25:
"A75 and A30;"

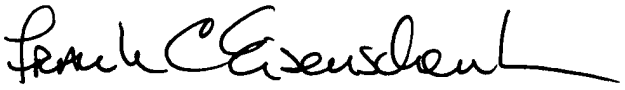
See Amendment dated 12/15/03, page 7, line 11, Claim 150:
--A75;--

True and correct copies of pages 7, 8, 10, 13, 17, 18, 24, 30, 31, 33, 37, 38, 40, 41, 42, 43, 44, 47, 62, 63, 64, 70, 72, 77, 82, 84, 91, 108, 115, 119, 127, 137, 147, 148, 149, 157, 161, 167, 169, 176, 195, 199, 204, and 206 of the specification as filed; Applicant's Amendments dated February 24, 2003 and December 15, 2003; and the Examiner's Amendment dated February 24, 2004 which support Applicants' assertion of errors on the part of the Patent Office accompanies this Certificate of Correction.

The Commissioner is authorized to charge the fee of $100.00 for the amendment to Deposit Account No. 19-0065. The Commissioner is also authorized to charge any additional fees as required under 37 CFR 1.20(a) to Deposit Account No. 19-0065. Two copies of this letter are enclosed for Deposit Account authorization.

Approval of the Certificate of Correction is respectfully requested.

Respectfully submitted,



Frank C. Eisenschenk, Ph.D.
Patent Attorney
Registration No. 45,332
Phone No.: 352-375-8100
Fax No.: 352-372-5800
Address: P.O. Box 142950
Gainesville, FL 32614-2950

FCE/mv
Attachments: Certificate of Correction
Copies of pages 7, 8, 10, 13, 17, 18, 24, 30, 31, 33, 37, 38, 40, 41, 42, 43, 44, 47, 62, 63, 64, 70, 72, 77, 82, 84, 91, 108, 115, 119, 127, 137, 147, 148, 149, 157, 161, 167, 169, 176, 195, 199, 204, and 206 of the specification as filed; the applicant's Amendments dated February 24, 2003 and December 15, 2003; and the Examiner's Amendment dated February 24, 2004
Two copies of this letter

Patent Reads:	**Application Reads:**
Column 4, Line 24: "and (consenus"	See Amendment dated 12/15/03, page 2: --and 22 (consensus)--
Column 4, Line 59: "82393. SEQ ID NO: 12"	See Examiner's Amendment dated 02/24/04, page 3: --82393.--
Column 5, Line 60: "Polynucleotides"	Page 7, Line 2: --polynucleotides--
Column 7, Line 17: "Volypeptide"	Page 8, Line 26: --polypeptide--
Column 8, Line 27: "heterozyposity"	Page 10, Line 6: --heterozygosity rate--
Column 11, Line 14: "bronchial En asthma"	Page 13, Lines 20-21: --bronchial asthma--
Column 14, Line 1: "BLASr"	Page 17, Line 8: --BLAST--
Column 15, Line 15: "exon 7, exon 9"	Page 18, Lines 29-30: --exon 7, exon 8, exon 9--
Column 20, Line 7: "nucleotide c at"	Page 24, Lines 30-31: --nucleotide C at--
Column 25, Line 8: "4400-8000"	Page 30, Line 24: --44001-48000--
Column 25, Line 44: "3'-regulatory,polynucleotide"	Page 31, Line 12: --3'-regulatory polynucleotide --

Column 27, Line 18: "seven let"	Page 33, Line 17: --seven *tet*--
Column 27, Line 34: "the let operator"	Page 33, Line 27: --the *tet* operator--
Column 31, Line 12: "24042406"	Page 38, Line 13: --2404-2406--
Column 31, Line 49: "nucleotide Tat positions"	Page 38, Line 35: --nucleotide T at positions--
Column 32, line 62: "thereto;"	Page 40, Lines 16-17: --thereto; and--
Column 33, Line 55: "A118, A123"	Page 41, Lines 20-21: --A118, and A123--
Column 34, Line 42: "A92A94to"	Page 42, Line 20: --A92, A94 to --
Column 35, Line 3: "A1107"	Page 43, Line 3: --A107--
Column 35, Line 9: "and complements"	Page 43, Lines 6-7: --and the complements--
Column 35, Line 13: "A55A56A59A92A94to"	Page 43, Line 9: --A55, A56, A59, A92, A94 to--
Column 35, Line 13: "A108A111 to A113A115"	Page 43, Line 9: --A108, A111 to A113, A115--
Column 36, Line 13: "^{32}p"	Page 44, Line 19: --^{32}P--
Column 38, Line 45: "VLSIPSυ"	Page 47, Line 25: --VLSIPS™--

Column 50, Line 16:
"A123 complements"

Page 62, Lines 10-11:
--A123 to A125, and the complements"--

Column 50, Line 20:
"A114, A122"

Page 62, Lines 13-14:
--A114, and A122--

Column 50, Line 25:
"to 113, A115, A117, and"

Page 62, Lines 16-17:
--to A113, A115 to A117, and--

Column 50, Line 64:
"A123 complements"

Page 63, Lines 6-7:
--A123 to A125, and the complements--

Column 51, Line 1:
"A109, A122,"

Page 63, Lines 9-10:
--A109, A110, A114, and A122--

Column 52, Line 10:
"are Q selected"

Page 64, Lines 22-23:
--are selected--

Column 56, Line 48:
"the w biallelic"

Page 70, Lines 15-16:
--the biallelic--

Column 58, Line 25:
"A109, A122"

Page 72, Lines 18-19:
--A109, A110, A114, and A122--

Column 58, Line 29:
"A94 to A01,"

Page 72, Line 21:
--A94 to A101,--

Column 62, Line 9:
"Eli"

Page 77, Line 14:
--E11--

Column 66, Line 25:
"about IS"

Page 82, Line 29:
--about 15--

Column 67, Line 27:
"WO 98120165"

Page 84, Line 4:
--WO 98/20165--

Column 73, Lines 22-23:
"with overlapping"

Page 91, Lines 21-22:
--with non-overlapping--

Column 87, Line 2: "(a_i,a_j, a_i,b_j: b_i,a_j)"	Page 108, Line 19: --(a_i,a_j; a_i,b_j; b_i,a_j)--
Column 87, Line 7: "θ4=—=frequency"	Page 108, Line 22: --θ4= - - = frequency--
Column 87, Line 9: "θ3=—+=frequency"	Page 108, Line 23: --θ3= - + = frequency--
Column 87, Line 11: "θ2=+—=frequency"	Page 108, Line 24: --θ2= + - = frequency--
Column 87, Line 14: "(M_i; M_j)"	Page 108, Line 26: --(M_i, M_j)--
Column 87, Line 21: "D_{aiaj} ($2n_1$+"	Page 108, Line 30: --D_{aiaj}= ($2n_1$+--
Column 92, Line 29: "AS5"	Page 115, Line 18: --A55--
Column 95, Lines 41-42: "preferred 1 5 embodiment"	Page 119, Lines 14-15: --preferred embodiment--
Column 101, Line 49: "be I,X;1 expressed"	Page 127, Lines 13-14: --be expressed--
Column 109, Line 22: "one-genes"	Page 137, Line 10: --*onc*-genes--
Column 117, Line 15: "(ISPR)"	Page 147, Line 13: --(SPR)--
Column 117, Line 17: "carboxymethyi"	Page 147, Line 14: --carboxymethyl--
Column 118, Line 16: "(MaTa gal4"	Page 148, Line 25: --(MATa gal4--

Column 118, Line 18: "lacZmet'),""	Page 148, Line 26: --lacZmet-)--
Column 118, Line 52: "GALA"	Line 149, Line 13: --GAL4--
Column 125, Line 6: "more CF particularly"	Page 157, Lines 22-23: --more particularly--
Column 127, Line 57: "monoleatc"	Page 161, Line 8: --monoleate--
Column 132, Line 42: "CAROM"	Page 167, Line 1: --CD-ROM--
Column 134, Line 47: "or a By polypeptide"	Page 169, Lines 20-21: --or a polypeptide--
Column 136, Line 47: "TTATAA Box"	See Amendment dated 02/24/03, page 2: --TAATAA Box--
Column 136, Line 48: "TATAA"	See Amendment dated 02/24/03, page 2: --TAATAA--
Column 140, Line 10: "[BLANK]"	Page 176, Line 6: --PCR assays were performed using the following protocol:--
Column 158, Line 1 of Table 6: "analysis of the"	Page 195, Line 1: --analysis for the--

Column 159, Line 2:
"analysis of the"

Page 195, Line 1:
--analysis for the--

Column 159, Line 6 of Table 6:
"

"

Page 195, Lines 3-5:
--

	A2	A30	A41	A55	A57	A75
frequency %	60/67 (A)	64/66 (T)	73/71 (T)	64/68 (C)	65/69 (G)	94/95 (G)
abs diff freq. all	-7,4	-2,0	2,7	-3,7	-3,9	-1
pvalue	7,7e-03	4,3e-01	2,9e-01	1,6e-01	1,4e-01	3,4e-01
Cases/controls ↓						

--

Column 164, Line 14:
"and b dose"

Page 199, Lines 20-21:
--and dose--

Column 164, Line 34:
"they arc"

Page 199, Line 35:
--they are--

Column 167, Line 25:
"Baltimore, 1991 Ouchterlony, O."

Page 204, Lines 30-31:
--Baltimore, 1991
Ouchterlony, O.--

Column 168, Line 22:
"119-1123"

Page 206, Line 6:
--1119-1123--

Column 168, Line 26:
"Stembert"

Page 206, Line 10:
--Sternberg--

Column 168, Line 26:
"5:397404"

Page 206, Line 10:
--5:397-404--

Column 168, Line 27:
"Strinmatter"

Page 206, Line 11:
--Strittmatter--

Patent Reads:

Application Should Read:

Column 569, Line 28:
"and b) performed"

See Amendment dated 12/15/03, page 6, line 6, Claim 137:
--and b) are performed--

Column 570, Line 25:
"A75 and A30;"

See Amendment dated 12/15/03, page 7, line 11, Claim 150:
--A75;--

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

7590 02/24/2004

SALIWANCHIK LLOYD & SALIWANCHIK
A PROFESSIONAL ASSOCIATION
2421 N.W. 41ST STREET
SUITE A-1
GAINESVILLE, FL 326066669

EXAMINER	
MAHATAN, CHANNING	
ART UNIT	PAPER NUMBER
1631	

DATE MAILED: 02/24/2004

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/326,402	06/04/1999	MARTA BLUMENFELD	GENSET.030A	4132

TITLE OF INVENTION: POLYMORPHIC MARKERS OF PROSTATE CARCINOMA TUMOR ANTIGEN-1(PCTA-1)

APPLN. TYPE	SMALL ENTITY	ISSUE FEE	PUBLICATION FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	NO	$1330	$0	$1330	05/24/2004

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE REFLECTS A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE APPLIED IN THIS APPLICATION. THE PTOL-85B (OR AN EQUIVALENT) MUST BE RETURNED WITHIN THIS PERIOD EVEN IF NO FEE IS DUE OR THE APPLICATION WILL BE REGARDED AS ABANDONED.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:
A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status is changed, pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above and notify the United States Patent and Trademark Office of the change in status, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check the box below and enclose the PUBLICATION FEE and 1/2 the ISSUE FEE shown above.

☐ Applicant claims SMALL ENTITY status. See 37 CFR 1.27.

II. PART B - FEE(S) TRANSMITTAL should be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). Even if the fee(s) have already been paid, Part B - Fee(s) Transmittal should be completed and returned. If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PTOL-85 (Rev. 11/03) Approved for use through 04/30/2004.

Notice of Allowability

Application No.	Applicant(s)	
09/326,402	BLUMENFELD ET AL.	
Examiner	**Art Unit**	
Channing S Mahatan	1631	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *19 December 2003*.

2. ☒ The allowed claim(s) is/are *121-137,140,141,150,165 and 166*.

3. ☒ The drawings filed on *04 March 2003* are accepted by the Examiner.

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

6. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

7. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO-1449 or PTO/SB/08), Paper No./Mail Date _____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application (PTO-152)
6. ☒ Interview Summary (PTO-413), Paper No./Mail Date *04 February 2004* .
7. ☒ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

DETAILED ACTION

APPLICANTS' ARGUMENTS

Applicants' arguments, filed 19 December 2003, have been fully considered and are deemed to be persuasive.

EXTENSION OF TIME

An extension of time under 37 C.F.R. § 1.136(a) is required in order to make an Examiner's amendment which places this application in condition for allowance. During a telephone conversation conducted on 04 February 2004, Frank C. Eisenschenk requested an extension of time for 2 MONTH(S) and authorized the Director to charge Deposit Account No. 19-0065 the required fee of $420.00 for this extension and authorized the following Examiner's amendment. Should the changes and/or additions be unacceptable to Applicants, an amendment may be filed as provided by 37 C.F.R. § 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

EXAMINER'S AMENDMENT

Authorization for this Examiner's amendment was given in a telephone interview with Frank C. Eisenschenk on 04 February 2004 and 05 February 2004.

Please amend the following claims:

121. (currently amended) A method of genotyping comprising determining the identity of a nucleotide at a biallelic marker of SEQ ID NO: 12, or the complement of said nucleotide, in a biological sample, wherein said biallelic marker is selected from the group consisting of biallelic markers of SEQ ID NO: 12 located at positions 402 (A2), 67092 (A30), 68525 (A41), 82234 (A55), 82393 (A57), and 87713 (A75).

135. (currently amended) A method of detecting an association between a haplotype and a trait, comprising the steps of:

a) estimating the frequency of at least one haplotype in a trait positive population ~~according to the method of claim 133~~;

b) estimating the frequency of said haplotype in a control population, wherein steps a) and b) are performed according to the method of claim 133;

c) determining whether a statistically significant association exists between said haplotype and said trait, wherein said trait is familial or sporadic prostate cancer.

Please substitute the following paragraph, beginning at page 5, line 26, with the following paragraph:

SEQ ID NO: 12 is a version of the polynucleotide sequence of the PCTA-1 gene of SEQ ID NO: 1 that provides symbols appropriate for indicated allelic substitutions at nucleotide positions 402, 67092, 68525, 82234, and 82393. ~~SEQ ID NO: 12~~

REASONS FOR ALLOWANCE

The following is an Examiner's statement of reasons for allowance:

The filing of the amendments to claims have overcome all grounds of rejection in the Office Action, mailed 21 May 2003, and no other grounds for rejection are present. Therefore, claims 121-137, 140, 141, 150, 165, and 166 are deemed allowable. No pending United States applications have been identified with claims directed to the same invention as claimed herein. Any comments considered necessary by Applicants must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee.

Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

EXAMINER INFORMATION

Papers related to this application may be submitted to Technical Center 1600 by facsimile transmission. Papers should be faxed to Technical Center 1600 via the PTO Fax Center located in Crystal Mall 1. The faxing of such papers must conform with the notices published in the Official Gazette, 1096 OG 30 (November 15, 1988), 1156 OG 61 (November 16, 1993), and 1157 OG 94 (December 28, 1993) (See 37 C.F.R. § 1.6(d)). The CM1 Fax Center number is either (703) 872-9306.

Any inquiry concerning this communication or earlier communications from the Examiner should be directed to Channing S. Mahatan whose telephone number is (571) 272-0717. The Examiner can normally be reached on M-F (8:30-5:00).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Michael P. Woodward, Ph.D., can be reached on (571) 272-0722.

Any inquiry of a general nature or relating to the status of this application should be directed to Legal Instruments Examiner, Tina M. Plunkett, whose telephone number is (571) 272-0549 or to the Technical Center receptionist whose telephone number is (703) 308-0196.

Date: February 5, 2004

Examiner Initials: CSM

Marianne P. Allen

MARIANNE P. ALLEN
PRIMARY EXAMINER
AU 1631

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450 Alexandria, VA 22313 on December 15, 2003.

Frank C. Eisenschenk

Frank C. Eisenschenk, Ph.D., Patent Attorney

AMENDMENT UNDER 37 C.F.R. §1.116
Examining Group 1631
Patent Application
Docket No. GEN-T112XC1
Serial No. 09/326,402

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Examiner	:	Channing S. Mahatan
Art Unit	:	1631
Applicants	:	Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov
Serial No.	:	09/326,402
Filed	:	June 4, 1999
Confirm. No.	:	4132
For	:	Polymorphic Markers of the Prostate Carcinoma Tumor Antigen-1 (PCTA-1)

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313

COPY

AMENDMENT UNDER 37 C.F.R. §1.116

Sir:

A Notice of Appeal was filed in this matter on October 21, 2003. Thus, Applicants respectfully submit that no fees or extensions of time are required by this paper. In response to the Office Action dated May 21, 2003, please amend the above-identified patent application as follows:

In the Specification

Please substitute the paragraph beginning on page 5, line 5, with the following paragraph:

Figures 7A-D is an alignment of the mouse and human PCTA-1 proteins. The amino acid sequences provided in the alignment are presented as SEQ ID NOs: 13 (leg2), 14 (leg1), 15 (PCTA), 16 (PCTA.var), 17 (PCTA.mus), 18 (gal9-1), 19 (gal), 20 (leg7), 21 (gal4), and 22 (consensus sequence).

Please substitute the following paragraph, beginning at page 5, line 26, with the following paragraph (added by Applicants' Amendment dated February 24, 2003):

SEQ ID NO:12 ~~contains a genomic~~is a version of the polynucleotide sequence of the *PCTA-1* gene of SEQ ID NO: 1 that provides symbols appropriate for indicated allelic substitutions at nucleotide positions 402, 67092, 68525, 82234, and 82393. SEQ ID NO: 12 ~~comprising the 5' regulatory region (upstream untranscribed region), the exons (0, 1, 2, 3, 4, 5, 6, 6bis, 7, 8, 9, 9bis, and 9ter) and introns, and the 3' regulatory region (downstream untranscribed region). This sequence is identical to that of SEQ ID NO:1 and these SEQ ID NOs may be used interchangeably throughout the subject application.~~

Please replace pages 1-157 of the Substitute Sequence Listing submitted to the Patent Office on February 24, 2003, with the accompanying Sequence Listing as new pages 1-212.

In the Claims

1-120. (canceled)

121. (currently amended) A method of genotyping comprising determining the identity of a nucleotide at a ~~PCTA-1-related~~ biallelic marker of SEQ ID NO: 12, or the complement of said nucleotide, in a biological sample, wherein said biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 402 (A2), 67092 (A30), 68525 (A41), 82234 (A55), 82393 (A57), and 87713 (A75).

122. (currently amended) ~~A~~ The method according to claim 121, wherein said biological sample is derived from a single subject.

123. (currently amended) ~~A~~ The method according to claim 122, wherein the identity of the nucleotides at said biallelic marker is determined for both copies of said biallelic marker present in said individual's genome.

124. (currently amended) ~~A~~ The method according to claim 121, wherein said biological sample is derived from multiple subjects.

125. (currently amended) ~~A~~ The method according to claim 121, further comprising amplifying a portion of said sequence comprising the biallelic marker prior to said determining step.

126. (currently amended) ~~A~~ The method according to claim 125, wherein said amplifying is performed by PCR.

127. (currently amended) ~~A~~ The method according to claim 121, wherein said determining is performed by a hybridization assay.

128. (currently amended) ~~A~~ The method according to claim 121, wherein said determining is performed by a sequencing assay.

129. (currently amended) ~~A~~ The method according to claim 121, wherein said determining is performed by a microsequencing assay.

130. (currently amended) ~~A~~ The method according to claim 121, wherein said determining is performed by an enzyme-based mismatch detection assay.

131. (currently amended) A method of estimating the frequency of an allele of a ~~PCTA-1-related~~ biallelic marker of SEQ ID NO: 12 in a population comprising:

a) genotyping individuals from said population for said biallelic marker according to the method of claim 121; and

b) determining the proportional representation of said biallelic marker in said population.

132. (currently amended) A method of detecting an association between a genotype and a trait, comprising the steps of:

a) performing the method of claim 131 to determine the frequency of at least one biallelic marker of SEQ ID NO: 12 in trait positive population;

b) performing the method of claim 131 to determine the frequency of at least one biallelic marker of SEQ ID NO: 12 in a control population; and

c) determining whether a statistically significant association exists between said genotype and said trait, wherein said trait is familial or sporadic prostate cancer.

~~A method of detecting an association between a genotype and a trait, comprising the steps of:~~

~~a) determining the frequency of at least one PCTA-1-related biallelic marker in trait positive population according to the method of claim 131;~~

~~b) determining the frequency of at least one PCTA-1-related biallelic marker in a control population according to the method of claim 131; and~~

~~c) determining whether a statistically significant association exists between said genotype and said trait.~~

133. (currently amended) A method of estimating the frequency of a haplotype for a set of biallelic markers of SEQ ID NO: 12 in a population, comprising:

a) genotyping at least one ~~PCTA-1-related~~ biallelic marker of SEQ ID NO: 12 according to claim 122 for each individual in said population;

b) genotyping a second biallelic marker of SEQ ID NO: 12 by determining the identity of the ~~nucleotides~~ nucleotide at said second biallelic marker for both copies of said second biallelic marker present in the genome of each individual in said population; and

c) applying a haplotype determination method to the identities of the nucleotides determined in steps a) and b) to obtain an estimate of said frequency.

134. (currently amended) ~~A~~ The method according to claim 133, wherein said haplotype determination method is selected from the group consisting of asymmetric PCR amplification, double PCR amplification of specific alleles, the Clark algorithm, or an expectation-maximization algorithm.

135. (currently amended) A method of detecting an association between a haplotype and a trait, comprising the steps of:

a) estimating the frequency of at least one haplotype in a trait positive population according to the method of claim 133;

b) estimating the frequency of said haplotype in a control population according to the method of claim 133; and

c) determining whether a statistically significant association exists between said haplotype and said trait, wherein said trait is familial or sporadic prostate cancer.

136. (currently amended) ~~A~~ The method according to claim 132, wherein said genotyping steps a) and b) are performed on a single pooled biological sample derived from each of said populations.

137. (currently amended) ~~A~~ The method according to claim 132, wherein said genotyping steps a) and b) performed separately on biological samples derived from each individual in said populations.

138-139. (canceled)

140. (previously added) The method according to claim 132, wherein said control population is a trait negative population.

141. (previously amended) The method according to claim 132, wherein said control population is a random population.

142-149. (canceled)

150. (currently amended) A method of determining whether an individual is at risk of developing familial or sporadic prostate cancer, comprising:

a) genotyping at least one ~~PCTA-1-related~~ biallelic marker of SEQ ID NO: 12 according to the method of claim 123; ~~and~~

b) determining if the individual has a biallelic marker or combination of biallelic markers that is associated with familial or sporadic prostate cancer; and

c~~b~~) correlating the result of step b) ~~a)~~ with a biallelic marker or one or more combinations of biallelic markers that are~~is~~ associated with a risk of developing familial or sporadic prostate cancer;

wherein:

for familial cases of prostate cancer said biallelic marker or combinations of biallelic markers are selected from the group consisting of: 1) A30 and A41; 2) A30 and A57; 3); A30 and A55; 4) A2 and A41; 5) A2 and A30; 6) A30 and A75; 7) A2 and A55; 8) A2, A30, and A41; 9) A2, A30, and A57; 10) A2, A30, and A55; 11) A30, A41, and A55; 12) A30, A57, and A75; 13) A30, A55, and A75; 14) A30, A41, and A75; 15) A30, A55, and A57; 16) A30, A41, and A57; 17) A2, A30, A55, and A57; 18) A2, A30, A57, and A75; 19) A2, A30, A55, and A75; 20) A2, A30, A41, and A57; 21) A2, A30, A41, and A55; 22) A2, A30, A41, and A75; 23) A30, A55, A57, and A75; 24) A30, A41, A55, and A75; 25) A30, A41, A55, and A57; 26) A30, A41, A57, and A75 and A30; and 27) A30; and

for sporadic cases of prostate cancer said combinations of biallelic markers are selected from the group consisting of: 1) A2 and A55; 2) A2 and A57; 3) A41 and A55; 4) A41 and A57; 5) A2 and A41; 6) A30 and A75; 7) A2, A41, and A55; 8) A2, A55, and A57; 9) A2, A41, and A57; 10) A41, A55, and A57; 11) A2, A55, and A75; 12) A30, A41, and A57; 13) A30, A41, and A55; 14) A2, A30, A41, and A57; 15) A2, A30, A41, and A55; 16) A2, A41, A55, and A57; 17) A2, A41, A55, and A75; and 18) A30, A41, A55, and A57.

151-164. (canceled)

165. (previously added) The method according to claim 135, wherein said control population is a trait negative population.

166. (previously added) The method according to claim 135, wherein said control population is a random population.

167-182. (canceled)

Remarks

Claims 89-182 were pending in the subject application. By way of this amendment, claims 121-137 and 150 have been amended and claims 1-120, 138, 139, 142-149, 151-164, and 167-182 have been canceled. Thus, claims 121-137, 140, 141, 150, 165, and 166 are before the Examiner for consideration. The undersigned avers that no new matter is introduced by this amendment and that support for the newly presented claims can be found in the application and claims as originally filed. Accordingly, entry and consideration of the amendments presented herein is respectfully requested. Favorable consideration of the pending claims is also respectfully requested.

As an initial matter, the Examiner has objected to the subject specification on the grounds that it fails to comply with the requirements of 37 C.F.R. 1.821(a)(1) and (a)(2) regarding Figures 7A-7D. Applicants have amended the specification to indicate that the relationship between the sequences of the Figure and those provided in the Sequence Listing (e.g., SEQ ID NOs: 14-22). Accordingly, reconsideration and withdrawal of the objection is respectfully requested.

The Office Action has objected to the specification under 35 U.S.C. § 132 on the basis that new matter has been introduced into the specification. In a previous Office Action, the Patent Office required the insertion of a description of SEQ ID NO: 12 and an indication of its relationship to SEQ ID NO: 1 in order to remove alleged confusion as to the two sequences. Pursuant to a request from the previous examiner in this matter, SEQ ID NO: 12 was originally presented to the Patent Office for the purposes of simplifying the search by indicating the appropriate symbols for nucleotides at various positions of SEQ ID NO: 1 that were to correspond to an elected invention. In order to clarify the current situation, Applicants submit herewith a Submission of Sequence Listing Under §1.821, including a replacement Sequence Listing on paper and in computer readable format, for the above-referenced patent application that indicates all the relevant features of SEQ ID NO: 12 (drawn from originally presented SEQ ID NO: 1). As also indicated in the Sequence Listing, SEQ ID NO: 12 contains appropriate symbols for the allele substitutions at nucleotide positions 402, 67092, 68525, 82234, and 82393. Applicants have also amended the Brief Description of the Sequences to attend to the new matter issue raised in the previous Office Action and believe that the issue has now been resolved. I hereby certify that the paper and computer readable copies contain the same information and that no new material is added by this submission. Entry and consideration of the

substitute Sequence Listing is respectfully requested. Accordingly, reconsideration and withdrawal of the rejection is respectfully requested; however, the Examiner is respectfully requested to contact the undersigned should additional clarification of this issue be required.

Claims 121-141, 150-152, and 163-182 were rejected under 35 U.S.C. §112, first paragraph, because the specification, while being enabling for a method of identifying a nucleotide (species position 67092 of the elected SEQ ID NO: 12) at a PCTA-1 biallelic marker (A30, allele T, p value = 0.033) in familial prostate cancer, biallelic marker combinations for familial cases of prostate cancer (Table 5; 26 combinations), and biallelic marker combinations for sporadic prostate cancer (Table 6; 18 combinations) does not reasonably provide enablement for the identification of other nucleotides at other PCTA-1 biallelic markers in familial prostate cancer (*i.e.*, A2, A41, *etc.*), sporadic prostate cancer (*i.e.*, A2, A30, *etc.*), or in all other combinations. In the interest of expediting prosecution, the claims have been amended and it is respectfully submitted that this issue is moot. Reconsideration and withdrawal of the objection is respectfully requested.

Claim 121, 131, 132, 133, 150, 151, 152, 167-172, 177, and 178, and all claims dependent therefrom, have been rejected because of the recitation "PCTA-1 related biallelic marker". It is respectfully submitted that the phrase is well-defined in the specification (for example, at page 11, lines 24-25); however, in the interest of expediting prosecution in this matter, Applicants have amended the claim to recite "a biallelic marker of SEQ ID NO: 12".

Claims 121, 140, 151, 167-172, 178, and all claims dependent therefrom have been rejected as being indefinite in the recitation of "the complement". The Office Action argues that it remains unclear as to the scope of the phrase "the complement". Applicants again respectfully traverse. As is clearly set forth in claim 121, the claim recites "or the complement of said nucleotide". In the context of the claim, it is respectfully submitted that one skilled in the art would recognize that the phrase is referring to a nucleotide found at a biallelic marker of SEQ ID NO: 12 as is recited within the claim. It is respectfully submitted that one skilled in the art would not have read the phrase to include the full length polynucleotide, rather the phrase would have been interpreted to include the complement of a biallelic nucleotide recited within the claim. Accordingly, reconsideration and withdrawal of the rejection is respectfully requested.

Claim 150 was rejected as being indefinite as it was alleged that the claim language failed to recite a final process step that agreed back to the preamble. Applicants have amended the claim and respectfully request reconsideration and withdrawal of the rejection.

Claims 140-141 have been rejected as being indefinite in the recitation of the phrase "control population" as it is argued that the claim from which these claims depend lacks same recitation. While the record in the possession of the undersigned indicates proper antecedent basis for the phrase, Applicants respectfully submit that this issue is now moot as the claim from which claims 140-141 depend has now been re-written to attend to this issue. Accordingly, reconsideration and withdrawal of the rejection is respectfully requested.

In view of the foregoing remarks and amendments to the claims, the applicants believe that the currently pending claims are in condition for allowance, and such action is respectfully requested.

It should be understood that the amendments presented herein have been made solely to expedite prosecution of the subject application to completion and should not be construed as an indication of Applicants' agreement with or acquiescence in the Examiner's position. Applicants expressly reserve the right to pursue the invention(s) disclosed in the subject application, including any subject matter canceled or not pursued during prosecution of the subject application, in a related application.

The Commissioner is hereby authorized to charge any fees under 37 C.F.R. §§ 1.16 or 1.17 as required by this paper to Deposit Account 19-0065.

The applicants invite the Examiner to call the undersigned if clarification is needed on any of this response, or if the Examiner believes a telephonic interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Frank C. Eisenschenk, Ph.D.
Patent Attorney
Registration No. 45,332
Phone No.: 352-375-8100
Fax No.: 352-372-5800
Address: Saliwanchik, Lloyd & Saliwanchik
A Professional Association
2421 NW 41st Street, Suite A-1
Gainesville, FL 32606-6669

FCE/sl
Attachment: New pages 1-212 (Sequence Listing)

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to the U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202, on February 24, 2003.

Frank C. Eisenschenk, Ph.D., Patent Attorney

AMENDMENT UNDER 37 C.F.R. § 1.111
Examining Group 1631
Patent Application
Docket No. GEN-T112XC1
Serial No. 09/326,402

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Examiner	:	Channing Mahatan
Art Unit	:	1631
Applicants	:	Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov
Serial No.	:	09/326,402
Filed	:	June 4, 1999
Confirm. No.	:	4132
For	:	Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)

BOX SEQUENCE
U.S. Patent and Trademark Office
P.O. Box 2327
Arlington, VA 22202

AMENDMENT UNDER 37 C.F.R. § 1.111

Sir:

A Petition and fee for a three-month Extension of Time through and including Monday, February 24, 2003, accompanies this Amendment.

In response to the Office Action dated August 23, 2002, please amend the above-identified application as follows:

In the Specification:

Please substitute the following paragraphs:

In the specification at page 5, line 26, please insert the following paragraph:

SEQ ID NO:12 contains a genomic sequence of *PCTA-1* comprising the 5' regulatory region (upstream untranscribed region), the exons (0, 1, 2, 3, 4, 5, 6, 6bis, 7, 8, 9, 9bis, and 9ter) and introns, and the 3' regulatory region (downstream untranscribed region). This sequence is identical to that of SEQ ID NO:1 and these SEQ ID NOs may be used interchangeably throughout the subject application.

Please replace the paragraph beginning at page 169, line 29 with the following paragraph:

Figure 6 is a flow diagram illustrating one embodiment of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a list of each feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An example of such a database is produced by the University of Wisconsin Genetics Computer Group (Worldwide Web address: gcg.com).

Please delete the paragraph found at page 190, lines 19-21.

Please substitute the Drawings (sheets 1-11) submitted herewith for the previously pending set of drawings (sheets 1-11).

In the Claims:

Please substitute the following claims:

121. (Amended) A method of genotyping comprising determining the identity of a nucleotide at a PCTA-1-related biallelic marker, or the complement of said nucleotide, in a biological sample.

141. (Thrice Amended) The method according to claim 132, wherein said control population is a random population.

150. (Amended) A method of determining whether an individual is at risk of developing prostate cancer, comprising:

a) genotyping at least one PCTA-1-related biallelic marker according to the method of claim 123; and

b) correlating the result of step a) with one or more biallelic marker that is associated with a risk of developing prostate cancer.

151. (Thrice Amended) The method according to claim 121, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complement of said biallelic marker.

152. (Fourth Amendment) The method according to claim 150, wherein said PCTA-1-related biallelic marker is at least one PCTA-1-related biallelic marker located at position 402, 67092, 68525, 82234, or 82393 of SEQ ID NO. 1, and any complement of said biallelic marker.

167. (Twice Amended) The method according to claim 131, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complement of said biallelic marker.

168. (Twice Amended) The method according to claim 132, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819,

96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complement of said biallelic marker.

169. (Twice Amended) The method according to claim 133, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937, and any complement of said biallelic marker.

170. (Twice Amended) The method according to claim 135, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complement of said biallelic marker.

171. (Twice Amended) The method according to claim 150, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complement of said biallelic marker.

177. (Twice Amended) The method according to claim 133, wherein said PCTA-1-related biallelic marker is at least one PCTA-1-related biallelic marker located at position 402, 67092, 68525, 82234, or 82393 of SEQ ID NO. 1, and any complement of said biallelic marker.

178. (Amended) The method according to claim 177, wherein said PCTA-1-related biallelic marker is a combination of more than one PCTA-1-related biallelic marker or any complement of said biallelic marker.

Remarks

Claims 89-182 are pending in the subject application. In response to a restriction requirement, claims 121-141, 150-152, and 163-182 were elected for examination on March 11, 2002. Certain of the claims have been amended for the purpose of expediting the patent application process in a manner consistent with the Patent and Trademark Office Patent Business Goals (PBG), 65 Fed. Reg. 54603 (September 8, 2000) and to advance prosecution and facilitate the business interests of Applicant(s). Support for these new claims and the amendments to the pending claims can be found throughout the subject specification, including, for example, at page 8, about line 16. Favorable consideration of the claims now presented, in view of the remarks and amendments set forth herein, is earnestly solicited.

The Office Action of August 23, 2002 has objected to the drawings and required corrections. Submitted herewith are corrected drawings for this matter.

The Office Action has also objected to the application on the grounds that certain of the Figures contain sequences not provided in the sequence listing that was filed with the above-referenced application. Applicants have submitted a substitute sequence listing with this response and believe that this issue is now moot.

The Office Action of August 23, 2002 also objected to the disclosure regarding the identical nature of SEQ ID NOs:1 and 12. Applicants have attended to this issue by indicating that these sequences are identical and that SEQ ID NO:1 and SEQ ID NO:12 can be used interchangeably throughout the subject specification. Additionally, any references to browser-executable code have been removed from the specification. Accordingly, withdrawal of the objection is respectfully requested.

The Office Action has also objected to the specification at page 190, lines 19-21, on the grounds that the paragraph structure is confusing and/or correlation to Tables 5 or 6. Applicants have deleted this paragraph and request withdrawal of the objection.

Claims 121-141, 150-152, and 163-182 have been rejected under 35 U.S.C. § 112, first paragraph, because the specification, while enabling for a method of identifying a nucleotide (allele T, marker A30) in cases of sporadic prostate cancer, does not reasonably provide enablement for the

identification of other nucleotides at other PCTA-1 biallelic markers in familial or sporadic prostate cancer. The Office Action further argues that the specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to use the invention commensurate in scope with these claims. The Office Action argues that the statistical and numerical information provided in Tables 5 and 6 fails to enable the claimed invention and that the data is statistically insignificant as relates to the identification of sporadic or familial cases of prostate cancer. Applicants disagree and respectfully traverse.

As set forth at pages 105-106 and 189-193 of the subject application, haplotype association analysis that combines the A2 marker with other markers disclosed in the specification does provide for statistically significant associations of the haplotypes with familial or sporadic prostate cancer. For example, as set forth at pages 105-106, a three-marker haplotype including markers A2, A30, and A41 (ATT alleles respectively) was shown to be significantly associated with familial prostate cancer (Table 8, the "ATT" haplotype has a p-value of 2.5×10^{-7} for the familial early onset prostate cancer (see Example 5)). A two-marker haplotype including markers A2 and A57 (99-1605/112) (TA alleles, respectively) was shown to be significantly associated with sporadic prostate cancer. As shown in Table 8, the "TA" haplotype present a p-value of 3.4×10^{-5} for the sporadic informative prostate cancer (see Example 5). A second two-marker haplotype including markers A2 and A55 (5-2/178) (TT alleles, respectively) was shown to be significantly associated with prostate cancer, preferably with a sporadic prostate cancer. As shown in Table 8, the "TT" haplotype present a p-value of 1×10^{-5} for the sporadic informative prostate cancer (see Example 5). Permutation tests clearly validated the statistical significance of the association between these haplotypes and the prostate cancer (see Example 5).

As set forth at pages 189-193, combination of the A2 marker (allele A) with A30 (allele T) and A41 (allele T) provides for a statistically significant association of the haplotype with familial cases of prostate cancer. The combination of A2 with additional biallelic markers, such as A55 and/or A57 also provides for statistically significant associations with sporadic cases of prostate cancer. In view of these teachings of the specification, it is respectfully submitted that the invention is properly enabled and that undue experimentation and further direction is not required in view of the teachings of the specification. Withdrawal of the rejection is respectfully requested.

Even assuming that the enablement rejection of record is properly asserted, Applicants respectfully submit that the arguments advanced in support of this enablement rejection are not applicable to a number of the claims (claims 131-141) set forth in this application and that the rejection, as applied to these claims, is improper. For example, these claims are drawn to methods of genotyping comprising determining the identity of a nucleotide at a PCTA-1 related biallelic marker, or the complement thereof, in a biological sample; estimating the frequency of an allele of a PCTA-1-related biallelic marker in a population; detecting the association between a genotype and a trait; or estimating the frequency of a haplotype for a set of biallelic markers in a population. Applicants respectfully submit that one skilled in the art would be able to practice this aspect of the claimed invention without undue experimentation and that the specification enables the invention, as claimed. Accordingly, reconsideration and withdrawal of the rejection is respectfully requested.

Claims 121-141, 150-152, and 163-182 have been rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter that was not describe in the specification in such a way as to convey to one skilled in the art that the inventors were in possession of the claimed invention at the time the application was filed. Specifically, the Office Action argues that the specification fails to provide adequate written description of the claimed invention in that the claims are directed to gene sequences, mutated fragment sequences, allelic variants, splice variants, and so forth that have no written description in the application as filed. Applicants respectfully submit that this rejection is moot in view of the amendments made to the claims and the arguments presented in the traversal of the rejections set forth under 35 U.S.C. § 112, second paragraph.

Claims 121-141, 150-152, and 163-182 have been rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter regarded as the invention.

Claims 121, 140, 151, 152, 167-172, 177, and 178 have been indicated as indefinite in the recitation of "the complement". The Office Action indicates that the term can be interpreted as a sequence that is the same length and be the full and exact complement of the recited SEQ ID NO. Alternative interpretations are also given in the Office Action. Applicants respectfully submit that the term is fully understandable in the context of the claimed invention and that one skilled in the art would recognize that the phrase "complement thereof" related to the nucleotides identified as the

biallelic markers of the elected sequence. However, Applicants have amended the claims to more distinctly identify that such is the case and respectfully request withdrawal of the rejection.

Claims 121, 131, 132, 133, 150, 151, 152, 167, 168, 169, 170, 171, 172, 177, and 178, and the claims dependent therefrom, have been rejected as indefinite in the recitation of "PCTA-1-related biallelic marker". The Office Action argues that some degree of relatedness is implied by such a recitation. Applicants respectfully submit that the acceptability of the claim language depends on whether one of ordinary skill in the art would understand what is claimed, in light of the specification. As set forth in the specification at page 11, lines 22-25, the term "PCTA-1-related biallelic marker" relates to a set of biallelic markers in linkage disequilibrium with the PCTA-1 gene. Accordingly, it is respectfully submitted that the claims would not be vague or indefinite to one skilled in the art, and withdrawal of the rejection is respectfully requested.

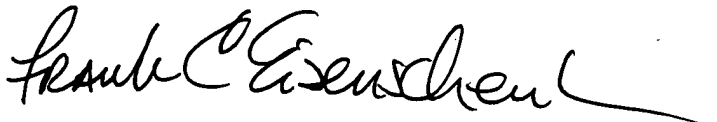
Claim 150 has been rejected as being indefinite the lack of a recitation of a final process step that relates back to the preamble of the claim. Applicants submit that this rejection is moot in view of the amendments made to the claim.

Claims 140 and 141 have been rejected as lacking antecedent basis in the claims from which each of the claims depend. With respect to the rejection as applied to claims 140 and 141, Applicants respectfully submit that the term "control population" can be found in step b) of claim 132. This rejection, as applied to "case control population" is also moot in view of the amendment made to the claims.

In view of the foregoing remarks and the amendments to the claims, the applicant believes that the pending claims are now in condition for allowance, and such action is respectfully requested. The Commissioner is hereby authorized to charge any fees under 37 C.F.R. §§ 1.16 or 1.17 as required by this paper to Deposit Account 19-0065.

Applicants also invite the Examiner to call the undersigned if clarification is needed on any of this response, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,



Frank C. Eisenschenk, Ph.D.
Patent Attorney
Registration No. 45,332
Phone: 352-375-8100
Fax No.: 352-372-5800
Address: Saliwanchik Lloyd & Saliwanchik
A Professional Association
2421 NW 41st Street, Suite A-1
Gainesville, FL 32606-6669

FCE/jaj

Attachments: Marked-Up Version of Amended Specification Paragraphs
Marked-Up Version of Amended Claims
Substitute Figures (Sheets 1-11)
Submission of Substitute Sequence List
Amendment regarding Sequence List
Paper Version of Substitute Sequence List
Electronic Version (CRF) of Substitute Sequence List
Petition for 3-Month Extension of Time (in duplicate)

Docket No. GEN-T112XC1
Serial No. 09/326,402

MARKED-UP VERSION OF AMENDED SPECIFICATION PARAGRAPH

Please replace the paragraph beginning at page 169, line 29 with the following paragraph:

Figure 6 is a flow diagram illustrating one embodiment of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a list of each feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An example of such a database is produced by the University of Wisconsin Genetics Computer Group (Worldwide Web address: ~~www.~~gcg.com).

MARKED-UP VERSION OF AMENDED CLAIMS

121. (Amended) A method of genotyping comprising determining the identity of a nucleotide at a PCTA-1-related biallelic marker, or the complement of said nucleotide~~thereof~~, in a biological sample.

141. (Thrice Amended) The method according to claim 132, wherein said ~~case~~-control population is a random population.

150. (Amended) A method of determining whether an individual is at risk of developing prostate cancer, comprising:

a) genotyping at least one PCTA-1-related biallelic marker according to the method of claim 123; and

b) correlating the result of step a) with one or more biallelic marker that is associated with a risk of developing prostate cancer.

151. (Thrice Amended) The method according to claim 121, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complements of said biallelic marker ~~thereof~~.

152. (Fourth Amendment) The method according to claim 150, wherein said PCTA-1-related biallelic marker is at least one PCTA-1-related biallelic marker located at position 402, 67092, 68525, 82234, or 82393 of SEQ ID NO. 1, and any complement of said biallelic marker ~~thereof~~.

167. (Twice Amended) The method according to claim 131, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complements of said biallelic marker ~~thereof~~.

168. (Twice Amended) The method according to claim 132, ~~,~~ wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819,

96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complements of said biallelic marker~~thereof~~.

169. (Twice Amended) The method according to claim 133, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937, and any complements of said biallelic marker~~thereof~~.

170. (Twice Amended) The method according to claim 135, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complements of said biallelic marker ~~thereof~~.

171. (Twice Amended) The method according to claim 150, wherein said PCTA-1-related biallelic marker is selected from the group consisting of biallelic markers of SEQ ID No:1 located at positions 278, 402, 472, 2955, 12167, 12536, 17593, 17606, 22079, 28964, 29003, 31077, 31766, 34791, 45751, 49847, 49855, 49886, 49900, 49906, 49921, 49939, 50256, 54955, 64239, 65436, 65496, 66967, 66987, 67092, 67129, 67229, 67433, 67723, 67834, 67955, 68213, 68263, 68375, 68477, 68525, 68594, 68610, 70566, 70728, 80038, 80118, 80170, 80183, 80435, 82090, 82165, 82169, 82218, 82234, 82268, 82393, 83587, 83643, 83644, 83808, 83881, 83884, 83909, 83937, 83947, 83982, 83988, 84047, 84092, 84145, 85202, 86259, 86323, 87713, 87735, 87787, 87806, 87860, 87902, 87980, 88012, 88215, 88283, 88297, 88442, 88471, 88480, 89394, 89464, 89471, 89658, 92760, 93023, 93055, 93247, 93319, 93323, 93388, 93393, 93418, 93515, 93726, 93903, 94170, 94218, 94269, 94290, 94422, 94432, 94720, 94989, 95261, 95340, 95497, 95770, 95819, 96145, 96181, 96350, 96951, 97144, 97276, 102267, 105937 and any complements of said biallelic marker ~~thereof~~.

177. (Twice Amended) The method according to claim 133, wherein said PCTA-1-related biallelic marker is at least one PCTA-1-related biallelic marker located at position 402, 67092, 68525, 82234, or 82393 of SEQ ID NO. 1, and any complement ~~thereof~~ of said biallelic marker.

178. (Amended) The method according to claim 177, wherein said PCTA-1-related biallelic marker is a combination of more than one PCTA-1-related biallelic marker or any complement of said biallelic marker~~thereof~~.

As used interchangeably herein, the terms "nucleic acids", "oligonucleotides", and polynucleotides" include RNA, DNA, or RNA/DNA hybrid sequences of more than one nucleotide in either single chain or duplex form. The term "nucleotide" as used herein as an adjective to describe molecules comprising RNA, DNA, or RNA/DNA hybrid sequences of any length in single-stranded or duplex form. The term "nucleotide" is also used herein as a noun to refer to individual nucleotides or varieties of nucleotides, meaning a molecule, or individual unit in a larger nucleic acid molecule, comprising a purine or pyrimidine, a ribose or deoxyribose sugar moiety, and a phosphate group, or phosphodiester linkage in the case of nucleotides within an oligonucleotide or polynucleotide. Although the term "nucleotide" is also used herein to encompass "modified nucleotides" which comprise at least one modifications (a) an alternative linking group, (b) an analogous form of purine, (c) an analogous form of pyrimidine, or (d) an analogous sugar, for examples of analogous linking groups, purine, pyrimidines, and sugars see for example PCT publication No. WO 95/04064. This may be especially oligonucleotides with α or β anomers, oligonucleotides with inter-nucleotide linkage of the phosphorothioate or methyl phosphonate type, or alternatively oligothionucleotide. The polynucleotide sequences of the invention may be prepared by any known method, including synthetic, recombinant, *ex vivo* generation, or a combination thereof, as well as utilizing any purification methods known in the art.

Throughout the present specification, the expression "nucleotide sequence" may be employed to designate indifferently a polynucleotide or a nucleic acid. More precisely, the expression "nucleotide sequence" encompasses the nucleic material itself and is thus not restricted to the sequence information (i.e. the succession of letters chosen among the four base letters) that biochemically characterizes a specific DNA or RNA molecule.

A "promoter" refers to a DNA sequence recognized by the synthetic machinery of the cell required to initiate the specific transcription of a gene.

A sequence which is "operably linked" to a regulatory sequence such as a promoter means that said regulatory element is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the nucleic acid of interest. As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence. More precisely, two DNA molecules (such as a polynucleotide containing a promoter region and a polynucleotide encoding a desired polypeptide or polynucleotide) are said to be "operably linked" if the nature of the linkage between the two polynucleotides does not (1) result in the introduction of a

frame-shift mutation or (2) interfere with the ability of the polynucleotide containing the promoter to direct the transcription of the coding polynucleotide.

The term "primer" denotes a specific oligonucleotide sequence which is complementary to a target nucleotide sequence and used to hybridize to the target nucleotide sequence. A primer serves as an initiation point for nucleotide polymerization catalyzed by either DNA polymerase, RNA polymerase or reverse transcriptase.

The term "probe" denotes a defined nucleic acid segment (or nucleotide analog segment, e.g., polynucleotide as defined herein) which can be used to identify a specific polynucleotide sequence present in samples, said nucleic acid segment comprising a nucleotide sequence complementary of the specific polynucleotide sequence to be identified.

The terms "base paired" and "Watson & Crick base paired" are used interchangeably herein to refer to nucleotides which can be hydrogen bonded to one another by virtue of their sequence identities in a manner like that found in double-helical DNA with thymine or uracil residues linked to adenine residues by two hydrogen bonds and cytosine and guanine residues linked by three hydrogen bonds (See Stryer, L., *Biochemistry,* 4th edition, 1995).

The terms "complementary" or "complement thereof" are used herein to refer to the sequences of polynucleotides which are capable of forming Watson & Crick base pairing with another specified polynucleotide throughout the entirety of the complementary region. For the purpose of the present invention, a first polynucleotide is deemed to be complementary to a second polynucleotide when each base in the first polynucleotide is paired with its complementary base. Complementary bases are, generally, A and T (or A and U), or C and G. "Complement" is used herein as a synonym of "complementary polynucleotide", "complementary nucleic acid" and "complementary nucleotide sequence". These terms are applied to pairs of polynucleotides based solely upon their sequences and not any particular set of conditions under which the two polynucleotides would actually bind.

The term "polypeptide" refers to a polymer of amino acids without regard to the length of the polymer; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not specify or exclude post-expression modifications of polypeptides, for example, polypeptides which include the covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like are expressly encompassed by the term polypeptide. Also included within the definition are polypeptides which contain one or more analogs of an amino acid (including, for example, non-naturally occurring amino acids, amino acids which only occur naturally in an unrelated biological system, modified amino acids from mammalian systems etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

Publication No. WO 84/03506, the disclosures of which are incorporated herein by reference in their entireties.

The term "allele" is used herein to refer to variants of a nucleotide sequence. A biallelic polymorphism has two forms. Diploid organisms may be homozygous or heterozygous for an allelic form.

The term "heterozygosity rate" is used herein to refer to the incidence of individuals in a population which are heterozygous at a particular allele. In a biallelic system, the heterozygosity rate is on average equal to $2P_a(1-P_a)$, where P_a is the frequency of the least common allele. In order to be useful in genetic studies, a genetic marker should have an adequate level of heterozygosity to allow a reasonable probability that a randomly selected person will be heterozygous.

The term "genotype" as used herein refers the identity of the alleles present in an individual or a sample. In the context of the present invention, a genotype preferably refers to the description of the biallelic marker alleles present in an individual or a sample. The term "genotyping" a sample or an individual for a biallelic marker consists of determining the specific allele or the specific nucleotide carried by an individual at a biallelic marker.

The term "mutation" as used herein refers to a difference in DNA sequence between or among different genomes or individuals which has a frequency below 1%.

The term "haplotype" refers to a combination of alleles present in an individual or a sample. In the context of the present invention, a haplotype preferably refers to a combination of biallelic marker alleles found in a given individual and which may be associated with a phenotype.

The term "polymorphism" as used herein refers to the occurrence of two or more alternative genomic sequences or alleles between or among different genomes or individuals. "Polymorphic" refers to the condition in which two or more variants of a specific genomic sequence can be found in a population. A "polymorphic site" is the locus at which the variation occurs. A single nucleotide polymorphism is a single base pair change. Typically a single nucleotide polymorphism is the replacement of one nucleotide by another nucleotide at the polymorphic site. Deletion of a single nucleotide or insertion of a single nucleotide, also give rise to single nucleotide polymorphisms. In the context of the present invention "single nucleotide polymorphism" preferably refers to a single nucleotide substitution. However, the polymorphism can also involve an insertion or a deletion of at least one nucleotide, preferably between 1 and 5 nucleotides. The nucleotide modification can also involve the presence of several adjacent single base polymorphisms. This type of nucleotide modification is usually called a "variable motif". Generally, a "variable motif" involves the presence of 2 to 10

The terms "an agent acting against prostate cancer" refers to any drug or compound that is capable of reducing the growth rate, rate of metastasis, or viability of tumor cells in a mammal, is capable of reducing the size or eliminating tumors in a mammal, or is capable of increasing the average life span of a mammal or human with cancer. Agents acting against prostate cancer also include compounds which are able to reduce the risk of cancer developing in a population, particularly a high risk population. Examples of agents acting against prostate cancer include hormonal therapeutic agents (for example, medroxyprogesterone acetate, estramustine phosphate, gonadotrophin releasing hormone (GnRH) agonists, anti-androgens such as flutamide, nilutamide, groserelin, and cyprosterone acetate, anti-gonadotropic agents such as stilboestrol and other oestrogenic agents, progestogens such as megestrol acetate) or chemotherapeutic agents (for example, carboplatin, cisplatin, methotrexate, mitomycin, epirubicin, vinblastine, 5-fluorouracyl, mitozantrone, cyclophosphamide, interferon, N-(4-hydroxyphenyl) retinamide (4HPR)). These agents can be used in combination.

The term "side effects to an agent acting against prostate cancer" refers to adverse effects of therapy resulting from extensions of the principal pharmacological action of the drug or to idiosyncratic adverse reactions resulting from an interaction of the drug with unique host factors. These side effects include, but are not limited to, adverse reactions such as dermatological, hematological or hepatological toxicities and further includes gastric and intestinal ulceration, disturbance in platelet function, renal injury, nephritis, vasomotor rhinitis with profuse watery secretions, angioneurotic edema, generalized urticaria, and bronchial asthma to laryngeal edema and bronchoconstriction, hypotension, sexual dysfunction, and shock. More particularly, the side effects can be nausea/vomiting, cardiovascular side effects such as deep vein thrombosis and fluid retention, and gynaecomastia.

The term "response to an agent acting against prostate cancer" refers to drug efficacy, including but not limited to ability to metabolize a compound, to the ability to convert a pro-drug to an active drug, and to the pharmacokinetics (absorption, distribution, elimination) and the pharmacodynamics (receptor-related) of a drug in an individual.

In the context of the present invention, a "positive response" to a medicament can be defined as comprising a reduction of the symptoms related to the disease, an increase of survival time or condition to be treated.

In the context of the present invention, a "negative response" to a medicament can be defined as comprising either a lack of positive response to the medicament which does not lead to a symptom reduction or an increase of survival time, or which leads to a side-effect observed following administration of the medicament.

positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Homology is evaluated using any of the variety of sequence comparison algorithms and programs known in the art. Such algorithms and programs include, but are by no means limited to, TBLASTN, BLASTP, FASTA, TFASTA, and CLUSTALW (Pearson and Lipman, 1988; Altschul et al., 1990; Thompson et al., 1994; Higgins et al., 1996; Altschul et al., 1993). In a particularly preferred embodiment, protein and nucleic acid sequence homologies are evaluated using the Basic Local Alignment Search Tool ("BLAST") which is well known in the art (see, e.g., Karlin and Altschul, 1990; Altschul et al., 1990, 1993, 1997). In particular, five specific BLAST programs are used to perform the following task:

(1) BLASTP and BLAST3 compare an amino acid query sequence against a protein sequence database;

(2) BLASTN compares a nucleotide query sequence against a nucleotide sequence database;

(3) BLASTX compares the six-frame conceptual translation products of a query nucleotide sequence (both strands) against a protein sequence database;

(4) TBLASTN compares a query protein sequence against a nucleotide sequence database translated in all six reading frames (both strands); and

(5) TBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.

The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (i.e., aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet et al., 1992; Henikoff and Henikoff, 1993). Less preferably, the PAM or PAM250 matrices may also be used (see, e.g., Schwartz and Dayhoff, eds., 1978). The BLAST programs evaluate the statistical significance of all high-scoring segment pairs identified, and preferably selects those segments which satisfy a user-specified threshold of significance, such as a user-specified percent homology. Preferably, the statistical significance of a high-scoring segment pair is evaluated using the statistical significance formula of Karlin (see, e.g., Karlin and Altschul, 1990).

Stringent Hybridization Conditions

By way of example and not limitation, procedures using conditions of high stringency are as follows: Prehybridization of filters containing DNA is carried out for 8 h to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 h at 65°C, the preferred hybridization temperature, in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10^6 cpm of ^{32}P-labeled probe. Alternatively, the hybridization step can be performed at 65°C in the presence of SSC buffer, 1 x SSC corresponding to 0.15M NaCl and 0.05 M Na citrate. Subsequently, filter washes can be done at 37°C for 1 h in a solution containing 2 x SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA, followed by a wash in 0.1 X SSC at 50°C for 45 min. Alternatively, filter washes can be performed in a solution containing 2 x SSC and 0.1% SDS, or 0.5 x SSC and 0.1% SDS, or 0.1 x SSC and 0.1% SDS at 68°C for 15 minute intervals. Following the wash steps, the hybridized probes are detectable by autoradiography. Other conditions of high stringency which may be used are well known in the art and as cited in Sambrook et al., 1989; and Ausubel et al., 1989. These hybridization conditions are suitable for a nucleic acid molecule of about 20 nucleotides in length. There is no need to say that the hybridization conditions described above are to be adapted according to the length of the desired nucleic acid, following techniques well known to the one skilled in the art. The suitable hybridization conditions may for example be adapted according to the teachings disclosed in the book of Hames and Higgins (1985) or in Sambrook et al.(1989).

Genomic Sequence Of The PCTA-1 Gene

The present invention relates to a purified and/or isolated nucleic acid corresponding to the genomic sequence of the *PCTA-1* gene. Preferably, this genomic *PCTA-1* sequence comprises the nucleotide sequence of SEQ ID No 1, a sequence complementary thereto, a fragment or a variant thereof.

The present invention encompasses the genomic sequence of *PCTA-1*. The *PTCA-1* gene sequence comprises a coding sequence including 13 exons included in SEQ ID No 1, namely exon 0, exon 1, exon 2, exon 3, exon 4, exon 5, exon 6, exon 6bis, exon 7, exon 8, exon 9, exon 9bis and exon 9ter, the intronic regions, the promoter, the 5'UTR, the 3'UTR, and regulatory regions located upstream and downstream of the coding region.

The localization of the exons and introns of the *PCTA-1* gene is detailed in Table A and is described as feature in SEQ ID No 1.

complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 708, 807, 1619, 1850, and 2126; a nucleotide C at positions 2062, 3505, and 3823; a nucleotide G at positions 709, 1971, 2059, 2060, 2061, 2107, 2140, and 3824; and a nucleotide T at positions 2533, and 2809 of SEQ ID No 3. Other preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises nucleotide positions selected from the group consisting of the nucleotide positions of SEQ ID No 3: 1355-1357, 1892-1894, 2090-2094, 2530-2532, 2619-2622, 3149-3152, 3183-3187, 3500-3501, 3508-3514, 3527-3532, 3533-3538, 3552-3557, 3746-3749, 3789-3792, 3797-3799, 3809-3815, 3912-3914 and 4056-4058. It should be noted that nucleic acid fragments of any size and sequence may also be comprised by the polynucleotides described in this section.

The inventors have also identified a species of *PCTA-1* cDNA comprising alternative exons to exon 9 which are called exons 9bis and 9ter. Its sequence is disclosed in SEQ ID No 4. The exon 9bis and 9ter correspond respectively to the beginning and the ends of the exon 9. The polynucleotide of the exon 9 located between exons 9bis and 9ter is spliced or deleted. The combination of exons 9bis and 9ter extends the ORF of the *PCTA-1* gene.

The main characteristics of this second *PCTA-1* cDNA comprising exons 0, 1, 2, 3, 4, 5, 6, 7, 8, 9bis and 9ter are detailed in Table B. The amino acid sequence of the new PCTA-1 protein encoded by this cDNA is disclosed in SEQ ID No 7. Particularly preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 4 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the nucleotide positions 1-162 of SEQ ID No 4. Further preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 4 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 253, 363, 527 and 2460 of SEQ ID No 4; a nucleotide C at position 1013 of SEQ ID No 4 and a nucleotide G at positions 176, and 749 of SEQ ID No 4. Additionally preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 4 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected

Thus, the present invention also concerns a purified or isolated nucleic acid comprising a polynucleotide which is selected from the group consisting of the 5' and 3' regulatory regions, or a sequence complementary thereto or a biologically active fragment or variant thereof. "5' regulatory region" refers to the nucleotide sequence located between positions 1 and 68647 of SEQ ID No 1. "3' regulatory region" refers to the nucleotide sequence located between positions 97156 and 106746 of SEQ ID No 1.

The invention also pertains to a purified or isolated nucleic acid comprising a polynucleotide having at least 95% nucleotide identity with a polynucleotide selected from the group consisting of the 5' and 3' regulatory regions, advantageously 99 % nucleotide identity, preferably 99.5% nucleotide identity and most preferably 99.8% nucleotide identity with a polynucleotide selected from the group consisting of the 5' and 3' regulatory regions, or a sequence complementary thereto or a biologically active fragment thereof.

Another object of the invention consists of purified, isolated or recombinant nucleic acids comprising a polynucleotide that hybridizes, under the stringent hybridization conditions defined herein, with a polynucleotide selected from the group consisting of the nucleotide sequences of the 5'- and 3' regulatory regions, or a sequence complementary thereto or a variant thereof or a biologically active fragment thereof.

Preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the following nucleotide positions of SEQ ID No 1: 1-4000, 4001-8000, 8001-12000, 12001-16000, 16001-20000, 20001-24000, 24001-28000, 28001-32000, 32001-36000, 36001-40000, 40001-44000, 44001-48000, 48001-52000, 52001-56000, 56001-60000, 60001-64000, 64001-68647. Particularly preferred nucleic acids of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the following nucleotide positions of SEQ ID No 1: 66647-68647.

"Biologically active" polynucleotide derivatives of SEQ ID No 1 are polynucleotides comprising or alternatively consisting of a fragment of said polynucleotide which is functional as a regulatory region for expressing a recombinant polypeptide or a recombinant polynucleotide in a recombinant cell host. It could act either as an enhancer or as a repressor.

For the purpose of the invention, a nucleic acid or polynucleotide is "functional" as a regulatory region for expressing a recombinant polypeptide or a recombinant polynucleotide if

said regulatory polynucleotide contains nucleotide sequences which contain transcriptional and translational regulatory information, and such sequences are "operably linked" to nucleotide sequences which encode the desired polypeptide or the desired polynucleotide.

The regulatory polynucleotides of the invention may be prepared from the nucleotide sequence of SEQ ID No 1 by cleavage using suitable restriction enzymes, as described for example in the book of Sambrook et al.(1989). The regulatory polynucleotides may also be prepared by digestion of SEQ ID No 1 by an exonuclease enzyme, such as Bal31 (Wabiko et al., 1986). These regulatory polynucleotides can also be prepared by nucleic acid chemical synthesis, as described elsewhere in the specification.

A preferred 5'-regulatory polynucleotide of the invention includes the 5'-untranslated region (5'-UTR) of the *PCTA-1* cDNA, or a biologically active fragment or variant thereof. A preferred 3'-regulatory polynucleotide of the invention includes the 3'-untranslated region (3'-UTR) of the *PCTA-1* cDNA, or a biologically active fragment or variant thereof.

A further object of the invention consists of a purified or isolated nucleic acid comprising:

a) a nucleic acid comprising a regulatory nucleotide sequence selected from the group consisting of:

(i) a nucleotide sequence comprising a polynucleotide of the 5' regulatory region or a complementary sequence thereto;

(ii) a nucleotide sequence comprising a polynucleotide having at least 95% of nucleotide identity with the nucleotide sequence of the 5' regulatory region or a complementary sequence thereto;

(iii) a nucleotide sequence comprising a polynucleotide that hybridizes under stringent hybridization conditions with the nucleotide sequence of the 5' regulatory region or a complementary sequence thereto; and

(iv) a biologically active fragment or variant of the polynucleotides in (i), (ii) and (iii);

b) a polynucleotide encoding a desired polypeptide or a nucleic acid of interest, operably linked to the nucleic acid defined in (a) above; and

c) Optionally, a nucleic acid comprising a 3'- regulatory polynucleotide, preferably a 3'- regulatory polynucleotide of the *PCTA-1* gene.

In a specific embodiment of the nucleic acid defined above, said nucleic acid includes the 5'-untranslated region (5'-UTR) of the *PCTA-1* cDNA, or a biologically active fragment or variant thereof. In a second specific embodiment of the nucleic acid defined above, said nucleic

this genomic sequence or cDNA harboring substitutions, deletions, or additions of one or more bases as regards to the *PCTA-1* nucleotide sequence of SEQ ID Nos 1, 2, 3, 4, 8, or a fragment thereof, these base substitutions, deletions or additions being located either in an exon, an intron or a regulatory sequence, but preferably in the 5'-regulatory sequence or in an exon of the *PCTA-1* genomic sequence or within a *PCTA-1* cDNA of SEQ ID Nos 2, 3, 4, or 8. In a preferred embodiment, the *PCTA-1* sequence comprises a biallelic marker of the present invention. In a preferred embodiment, the *PCTA-1* sequence comprises a biallelic marker of the present invention, preferably one of the biallelic markers A1 to A125 and the complements thereof.

The present invention embodies recombinant vectors comprising any one of the polynucleotides described in the present invention. More particularly, the polynucleotide constructs according to the present invention can comprise any of the polynucleotides described in the "*PCTA-1* cDNA Sequences" section, the "Coding Regions" section, and the "Oligonucleotide Probes And Primers" section.

A first preferred DNA construct is based on the tetracycline resistance operon *tet* from *E. coli* transposon Tn10 for controlling the *PCTA-1* gene expression, such as described by Gossen et al.(1992, 1995) and Furth et al.(1994). Such a DNA construct contains seven *tet* operator sequences from Tn10 (*tet*op) that are fused to either a minimal promoter or a 5'-regulatory sequence of the *PCTA-1* gene, said minimal promoter or said *PCTA-1* regulatory sequence being operably linked to a polynucleotide of interest that codes either for a sense or an antisense oligonucleotide or for a polypeptide, including a *PCTA-1* polypeptide or a peptide fragment thereof. This DNA construct is functional as a conditional expression system for the nucleotide sequence of interest when the same cell also comprises a nucleotide sequence coding for either the wild type (tTA) or the mutant (rTA) repressor fused to the activating domain of viral protein VP16 of herpes simplex virus, placed under the control of a promoter, such as the HCMVIE1 enhancer/promoter or the MMTV-LTR. Indeed, a preferred DNA construct of the invention comprise both the polynucleotide containing the *tet* operator sequences and the polynucleotide containing a sequence coding for the tTA or the rTA repressor.

In a specific embodiment, the conditional expression DNA construct contains the sequence encoding the mutant tetracycline repressor rTA, the expression of the polynucleotide of interest is silent in the absence of tetracycline and induced in its presence.

DNA Constructs Allowing Homologous Recombination: Replacement Vectors

A second preferred DNA construct will comprise, from 5'-end to 3'-end: (a) a first nucleotide sequence that is comprised in the *PCTA-1* genomic sequence; (b) a nucleotide

150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide G at positions 70728, 87860, 88297, 94432, and 95340 of SEQ ID No 1; a nucleotide A at positions 82218, 83644, 83808, 87787, 87806, 94218, and 97144 of SEQ ID No 1; a nucleotide C at positions 87902, 88215, 88283, 92760, 93726, and 94422 of SEQ ID No 1; and a nucleotide T at positions 93903, and 94170 of SEQ ID No 1. Other preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide G at positions 86435, 93592, 93680, 93681, 93682, 93728, 93761, and 95445 of SEQ ID No 1; a nucleotide A at positions 86434, 88355, 93240, 93471, and 93747of SEQ ID No 1; a nucleotide C at positions 93683, 95126, and 95444 of SEQ ID No 1; and a nucleotide T at positions 94154, and 94430 of SEQ ID No 1. Other preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises nucleotide positions selected from the group consisting of the nucleotide positions of SEQ ID No 1: 92975-92977, 93711-93715, 94151-94153, 94240-94243, 94770-94773, 94804-94808, 95121-95122, 95129-95135, 95148-95153, 95154-95159, 95173-95178, 95367-95374, 95410-95413, 95418-95420, 95430-95436, 95533-95535, and 95677-95677.

Another object of the invention is a purified, isolated, or recombinant polynucleotide comprising the nucleotide sequence of SEQ ID No 2, complementary sequences thereto, as well as allelic variants, and fragments thereof. Particularly preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 2 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the nucleotide positions 1-162 of SEQ ID No 2. Additional preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 2 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 253, 363, 527, 2471, and 5397 of SEQ ID No 2; a nucleotide C at positions 1013, 1979, and 2675 of SEQ ID No 2; a nucleotide G at positions 176, 749, 2685, 3593 of SEQ ID No 2; and a nucleotide T at positions 2156, and 2423

of SEQ ID No 2. Particularly preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 2 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 708, 807, 1493, 1724, and 2000; a nucleotide C at positions 1936, 3379, and 3697; a nucleotide G at positions 709, 1845, 1933, 1934, 1935, 1981, 2014, and 3698; and a nucleotide T at positions 2407, and 2683 of SEQ ID No 2. Other preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 2 or the complements thereof, wherein said contiguous span comprises nucleotide positions selected from the group consisting of the nucleotide positions of SEQ ID No 2: 1229-1231, 1964-1968, 2404-2406, 2493-2496, 3023-3026, 3057-3061, 3374-3375, 3382-3388, 3401-3406, 3407-3412, 3426-3431, 3620-3627, 3663-3666, 3671-3673, 3683-3689, 3786-3788 and 3930-3932.

A further object of the invention is a purified, isolated, or recombinant polynucleotide comprising the nucleotide sequence of SEQ ID No 3, complementary sequences thereto, as well as allelic variants, and fragments thereof. Particularly preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the following nucleotide positions of SEQ ID No 3: 1-162 and 747-872. Additional preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 253, 363, 527, 2597, and 5523 of SEQ ID No 3; a nucleotide C at positions 1139, 2105, and 2801 of SEQ ID No 3; a nucleotide G at positions 176, 875, 2811, 3719 of SEQ ID No 3; and a nucleotide T at positions 2282, and 2549 of SEQ ID No 3. Additional preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 3 or the complements thereof, wherein said contiguous span comprises at least one nucleotide selected from the group consisting of a nucleotide A at positions 708, 807, 1619, 1850, and 2126; a nucleotide C at positions 2062, 3505, and 3823; a nucleotide G at positions 709, 1971, 2059, 2060, 2061, 2107, 2140, and 3824; and a nucleotide T at positions 2533, and

least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of SEQ ID No 8 or the complements thereof, wherein said contiguous span comprises at least 1, 2, 3, 5, or 10 of the following nucleotide positions of SEQ ID No 8: 1-500, 501-1000, 1001-1500, and 1501-1738.

Thus, the invention also relates to nucleic acid probes characterized in that they hybridize specifically, under the stringent hybridization conditions defined above, with a nucleic acid selected from the group consisting of the nucleotide sequences:

a) 1-70715, 70795-82207, 82297-83612, 83824-85297, 85418-86388, 86446-87495, 87523-88294, 88384-89483, 89650-92748, 97156-98309, 98476-99329, 99491-100026, 100212-100281, 100396-100538, 100682-100833, 100995-101920, 102087-102970, 103264-103724, and 103753-106746 of SEQ ID No 1 or a variant thereof or a sequence complementary thereto;

b) 1-162 of SEQ ID No 2 or a variant thereof or a sequence complementary thereto;

c) 1-162 and 747-872 of SEQ ID No 3 or a variant thereof or a sequence complementary thereto;

d) 1-162 of SEQ ID No 4 or a variant thereof or a sequence complementary thereto; and

e) SEQ ID No 8 or a variant thereof or a sequence complementary thereto.

In a preferred embodiment, the oligonucleotides of the invention can hybridize with at least a portion of an intron or of the regulatory sequences of the *PCTA-1* gene. Particularly preferred oligonucleotides of the invention hybridize with a sequence comprised in an intron or in the regulatory sequences of the *PCTA-1* gene. In an other preferred embodiment, the oligonucleotides of the invention can hybridize with at least a portion of an exon selected from the group of exons 0, 1, 6bis, 9, and 9ter.

The present invention also concerns oligonucleotides and groups of oligonucleotides for the detection of alleles of biallelic markers of the *PCTA-1* gene, preferably those associated with cancer, preferably with prostate cancer, with an early onset of prostate cancer, with a susceptibility to prostate cancer, with the level of aggressiveness of prostate cancer tumors, with a modified or forthcoming expression of the *PCTA-1* gene, with a modified or forthcoming production of the PCTA-1 protein, or with the production of a modified PCTA-1 protein. These oligonucleotides are characterized in that they can hybridize with a *PCTA-1* gene, preferably with a polymorphic *PCTA-1* gene and more preferably with a region of a *PCTA-1* gene comprising a polymorphic site containing a specific allele associated with prostate cancer, with the level of aggressiveness of prostate cancer tumors or with modifications in the regulation of expression of the *PCTA-1* gene. These oligonucleotides are useful either as primers for use in

various processes such as DNA amplification and microsequencing or as probes for DNA recognition in hybridization analyses.

Therefore, another preferred embodiment of a probe according to the invention consists of a nucleic acid comprising a biallelic marker selected from the group consisting of A1 to A125 or the complements thereof, for which the respective locations in the sequence listing are provided in Table 2. In some embodiments, the oligonucleotides comprise the polymorphic base of a sequence selected from P1 to P125, and the complementary sequences thereto. In other embodiments, the oligonucleotides have a 3' terminus immediately adjacent to a polymorphic base in the *PCTA-1* gene, such as a polymorphic base comprised in one of the sequences P1 to P125, and the complementary sequence thereto. In other embodiments, the oligonucleotide is capable of discriminating between different alleles of a biallelic marker in the *PCTA-1* gene, including the biallelic markers A1 to A125 and the complements thereof.

In one embodiment the invention encompasses isolated, purified, and recombinant polynucleotides consisting of, or consisting essentially of a contiguous span of 8 to 50 nucleotides of any one of SEQ ID Nos 1, 2, 3, 4 and the complement thereof, wherein said span includes a *PCTA-1*-related biallelic marker in said sequence; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said contiguous span is 18 to 47 nucleotides in length and said biallelic marker is within 4 nucleotides of the center of said polynucleotide; optionally, wherein said polynucleotide consists of said contiguous span and said contiguous span is 25 nucleotides in length and said biallelic marker is at the center of said polynucleotide; optionally, wherein said polynucleotide consists of said contiguous span and said contiguous span is 47 nucleotides in length and said biallelic marker is at the center of said polynucleotide; optionally, wherein the 3' end of said contiguous span is present at the 3' end of said polynucleotide; and optionally, wherein the 3' end of said

contiguous span is located at the 3' end of said polynucleotide and said biallelic marker is present at the 3' end of said polynucleotide. In a preferred embodiment, said probes comprises, consists of, or consists essentially of a sequence selected from the following sequences: P1 to P125 and the complementary sequences thereto.

In another embodiment the invention encompasses isolated, purified and recombinant polynucleotides comprising, consisting of, or consisting essentially of a contiguous span of 8 to 50 nucleotides of SEQ ID Nos 1, 2, 3, 4, or the complements thereof, wherein the 3' end of said contiguous span is located at the 3' end of said polynucleotide, and wherein the 3' end of said polynucleotide is located within 20 nucleotides upstream of a *PCTA-1*-related biallelic marker in said sequence; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein the 3' end of said polynucleotide is located 1 nucleotide upstream of said *PCTA-1*-related biallelic marker in said sequence; and optionally, wherein said polynucleotide consists essentially of a sequence selected from the following sequences: D1 to D125 and E1 to E125.

In a further embodiment, the invention encompasses isolated, purified, or recombinant polynucleotides comprising, consisting of, or consisting essentially of a sequence selected from the following sequences: B1 to B47 and C1 to C47.

In an additional embodiment, the invention encompasses polynucleotides for use in hybridization assays, sequencing assays, and enzyme-based mismatch detection assays for determining the identity of the nucleotide at a *PCTA-1*-related biallelic marker in SEQ ID Nos 1, 2, 3, 4, or the complements thereof, as well as polynucleotides for use in amplifying segments of nucleotides comprising a *PCTA-1*-related biallelic marker in SEQ ID Nos 1, 2, 3, 4, or the complements thereof; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic

markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith.

The formation of stable hybrids depends on the melting temperature (Tm) of the DNA. The Tm depends on the length of the primer or probe, the ionic strength of the solution and the G+C content. The higher the G+C content of the primer or probe, the higher is the melting temperature because G:C pairs are held by three H bonds whereas A:T pairs have only two. The GC content in the probes of the invention usually ranges between 10 and 75 %, preferably between 35 and 60 %, and more preferably between 40 and 55 %.

A probe or a primer according to the invention has between 8 and 1000 nucleotides in length, or is specified to be at least 12, 15, 18, 20, 25, 35, 40, 50, 60, 70, 80, 100, 250, 500 or 1000 nucleotides in length. More particularly, the length of these probes and primers can range from 8, 10, 15, 20, or 30 to 100 nucleotides, preferably from 10 to 50, more preferably from 15 to 30 nucleotides. Shorter probes and primers tend to lack specificity for a target nucleic acid sequence and generally require cooler temperatures to form sufficiently stable hybrid complexes with the template. Longer probes and primers are expensive to produce and can sometimes self-hybridize to form hairpin structures. The appropriate length for primers and probes under a particular set of assay conditions may be empirically determined by one of skill in the art. A preferred probe or primer consists of a nucleic acid comprising a polynucleotide selected from the group of the nucleotide sequences of P1 to P125 and the complementary sequence thereto, B1 to B47, C1 to C47, D1 to D125, E1 to E125, for which the respective locations in the sequence listing are provided in Tables 1, 2, 3 and 4.

The primers and probes can be prepared by any suitable method, including, for example, cloning and restriction of appropriate sequences and direct chemical synthesis by a method such as the phosphodiester method of Narang et al.(1979), the phosphodiester method of Brown et al.(1979), the diethylphosphoramidite method of Beaucage et al.(1981) and the

solid support method described in EP 0 707 592, the disclosure of which is incorporated herein by reference in its entirety.

Detection probes are generally nucleic acid sequences or uncharged nucleic acid analogs such as, for example peptide nucleic acids which are disclosed in International Patent Application WO 92/20702, morpholino analogs which are described in U.S. Patents Numbered 5,185,444; 5,034,506 and 5,142,047, the disclosures of which are incorporated herein by reference in their entireties. The probe may have to be rendered "non-extendable" in that additional dNTPs cannot be added to the probe. In and of themselves analogs usually are non-extendable and nucleic acid probes can be rendered non-extendable by modifying the 3' end of the probe such that the hydroxyl group is no longer capable of participating in elongation. For example, the 3' end of the probe can be functionalized with the capture or detection label to thereby consume or otherwise block the hydroxyl group. Alternatively, the 3' hydroxyl group simply can be cleaved, replaced or modified, U.S. Patent Application Serial No. 07/049,061 filed April 19, 1993 describes modifications, which can be used to render a probe non-extendable.

Any of the polynucleotides of the present invention can be labeled, if desired, by incorporating any label known in the art to be detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include radioactive substances (including, ^{32}P, ^{35}S, ^{3}H, ^{125}I), fluorescent dyes (including, 5-bromodesoxyuridin, fluorescein, acetylaminofluorene, digoxigenin) or biotin. Preferably, polynucleotides are labeled at their 3' and 5' ends. Examples of non-radioactive labeling of nucleic acid fragments are described in the French patent No. FR-7810975 or by Urdea et al (1988) or Sanchez-Pescador et al (1988), the disclosures of which are incorporated herein by reference in their entireties. In addition, the probes according to the present invention may have structural characteristics such that they allow the signal amplification, such structural characteristics being, for example, branched DNA probes as those described by Urdea et al. in 1991 or in the European patent No. EP 0 225 807 (Chiron), the disclosures of which are incorporated herein by reference in their entireties.

A label can also be used to capture the primer, so as to facilitate the immobilization of either the primer or a primer extension product, such as amplified DNA, on a solid support. A capture label is attached to the primers or probes and can be a specific binding member which forms a binding pair with the solid's phase reagent's specific binding member (e.g. biotin and streptavidin). Therefore depending upon the type of label carried by a polynucleotide or a probe, it may be employed to capture or to detect the target DNA. Further, it will be understood that the polynucleotides, primers or probes provided herein, may, themselves, serve as the

Oligonucleotide Arrays

A substrate comprising a plurality of oligonucleotide primers or probes of the invention may be used either for detecting or amplifying targeted sequences in the *PCTA-1* gene and may also be used for detecting mutations in the coding or in the non-coding sequences of the *PCTA-1* gene.

Any polynucleotide provided herein may be attached in overlapping areas or at random locations on the solid support. Alternatively the polynucleotides of the invention may be attached in an ordered array wherein each polynucleotide is attached to a distinct region of the solid support which does not overlap with the attachment site of any other polynucleotide. Preferably, such an ordered array of polynucleotides is designed to be "addressable" where the distinct locations are recorded and can be accessed as part of an assay procedure. Addressable polynucleotide arrays typically comprise a plurality of different oligonucleotide probes that are coupled to a surface of a substrate in different known locations. The knowledge of the precise location of each polynucleotides location makes these "addressable" arrays particularly useful in hybridization assays. Any addressable array technology known in the art can be employed with the polynucleotides of the invention. One particular embodiment of these polynucleotide arrays is known as the Genechips™, and has been generally described in US Patent 5,143,854; PCT publications WO 90/15070 and 92/10092, the disclosures of which are incorporated herein by reference in their entireties. These arrays may generally be produced using mechanical synthesis methods or light directed synthesis methods which incorporate a combination of photolithographic methods and solid phase oligonucleotide synthesis (Fodor et al., 1991). The immobilization of arrays of oligonucleotides on solid supports has been rendered possible by the development of a technology generally identified as "Very Large Scale Immobilized Polymer Synthesis" (VLSIPS™) in which, typically, probes are immobilized in a high density array on a solid surface of a chip. Examples of VLSIPS™ technologies are provided in US Patents 5,143,854; and 5,412,087 and in PCT Publications WO 90/15070, WO 92/10092 and WO 95/11995, the disclosures of which are incorporated herein by reference in their entireties, which describe methods for forming oligonucleotide arrays through techniques such as light-directed synthesis techniques. In designing strategies aimed at providing arrays of nucleotides immobilized on solid supports, further presentation strategies were developed to order and display the oligonucleotide arrays on the chips in an attempt to maximize hybridization patterns and sequence information. Examples of such presentation strategies are disclosed in PCT Publications WO 94/12305, WO 94/11530, WO 97/29212 and WO 97/31256.

In another embodiment of the oligonucleotide arrays of the invention, an oligonucleotide probe matrix may advantageously be used to detect mutations occurring in the

The invention also encompasses the use of any polynucleotide for, or any polynucleotide for use in, determining the identity of one or more nucleotides at a *PCTA-1*-related biallelic marker. In addition, the polynucleotides of the invention for use in determining the identity of one or more nucleotides at a *PCTA-1*-related biallelic marker encompass polynucleotides with any further limitation described in this disclosure, or those following, specified alone or in any combination. Optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; Optionally, said polynucleotide may comprise a sequence disclosed in the present specification; Optionally, said polynucleotide may consist of, or consist essentially of any polynucleotide described in the present specification; Optionally, said determining may be performed in a hybridization assay, a sequencing assay, a microsequencing assay, or an enzyme-based mismatch detection assay; A preferred polynucleotide may be used in a hybridization assay for determining the identity of the nucleotide at a *PCTA-1*-related biallelic marker. Another preferred polynucleotide may be used in a sequencing or microsequencing assay for determining the identity of the nucleotide at a *PCTA-1*-related biallelic marker. A third preferred polynucleotide may be used in an enzyme-based mismatch detection assay for determining the identity of the nucleotide at a *PCTA-1*-related biallelic marker. A fourth preferred polynucleotide may be used in amplifying a segment of polynucleotides comprising a *PCTA-1*-related biallelic marker. Optionally, any of the polynucleotides described above may be attached to a solid support, array, or addressable array; Optionally, said polynucleotide may be labeled.

Additionally, the invention encompasses the use of any polynucleotide for, or any polynucleotide for use in, amplifying a segment of nucleotides comprising a *PCTA-1*-related biallelic marker. In addition, the polynucleotides of the invention for use in amplifying a segment of nucleotides comprising a *PCTA-1*-related biallelic marker encompass

polynucleotides with any further limitation described in this disclosure, or those following, specified alone or in any combination: Optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; Optionally, said polynucleotide may comprise a sequence disclosed in the present specification; Optionally, said polynucleotide may consist of, or consist essentially of any polynucleotide described in the present specification; Optionally, said amplifying may be performed by a PCR or LCR. Optionally, said polynucleotide may be attached to a solid support, array, or addressable array. Optionally, said polynucleotide may be labeled.

The primers for amplification or sequencing reaction of a polynucleotide comprising a biallelic marker of the invention may be designed from the disclosed sequences for any method known in the art. A preferred set of primers are fashioned such that the 3' end of the contiguous span of identity with a sequence selected from the group consisting of SEQ ID Nos 1, 2, 3, 4 or a sequence complementary thereto or a variant thereof is present at the 3' end of the primer. Such a configuration allows the 3' end of the primer to hybridize to a selected nucleic acid sequence and dramatically increases the efficiency of the primer for amplification or sequencing reactions. Allele specific primers may be designed such that a polymorphic base of a biallelic marker is at the 3' end of the contiguous span and the contiguous span is present at the 3' end of the primer. Such allele specific primers tend to selectively prime an amplification or sequencing reaction so long as they are used with a nucleic acid sample that contains one of the two alleles present at a biallelic marker. The 3' end of the primer of the invention may be located within or at least 2, 4, 6, 8, 10, 12, 15, 18, 20, 25, 50, 100, 250, 500, or 1000 nucleotides upstream of a *PCTA-1*-related biallelic marker in said sequence or at any other location which is appropriate for their intended use in sequencing, amplification or the location of novel sequences or markers. Thus, another set of preferred amplification primers comprise an isolated

polynucleotide consisting essentially of a contiguous span of 8 to 50 nucleotides in a sequence selected from the group consisting of SEQ ID Nos 1, 2, 3, 4 or a sequence complementary thereto or a variant thereof, wherein the 3' end of said contiguous span is located at the 3'end of said polynucleotide, and wherein the 3'end of said polynucleotide is located upstream of a *PCTA-1*-related biallelic marker in said sequence. Preferably, those amplification primers comprise a sequence selected from the group consisting of the sequences B1 to B47 and C1 to C47. Primers with their 3' ends located 1 nucleotide upstream of a biallelic marker of *PCTA-1* have a special utility as microsequencing assays. Preferred microsequencing primers are described in Table 4. Optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; Optionally, microsequencing primers are selected from the group consisting of the nucleotide sequences D1 to D125 and E1 to E125. More preferred microsequencing primers are selected from the group consisting of the nucleotides sequences D15, D24, D30, D34, D36, D38, D41, D44, D50, D53, D54, D56, D57, D59, D76, D85, D93, D108, D111, D115, D124, E11, E14, E22, E25, E26, E35, E42, E52, E53, E55, E56, E60, E61, E64, E73, E75, E93, E96.

The probes of the present invention may be designed from the disclosed sequences for any method known in the art, particularly methods which allow for testing if a marker disclosed herein is present. A preferred set of probes may be designed for use in the hybridization assays of the invention in any manner known in the art such that they selectively bind to one allele of a biallelic marker, but not the other under any particular set of assay conditions. Preferred hybridization probes comprise the polymorphic base of either allele 1 or allele 2 of the considered biallelic marker. Optionally, said biallelic marker may be within 6, 5, 4, 3, 2, or 1 nucleotides of the center of the hybridization probe or at the center of said probe. In a preferred embodiment, the probes are selected from the group consisting of the sequences P1 to P125 and the complementary sequence thereto.

Preferably, the amplified DNA is subjected to automated dideoxy terminator sequencing reactions using a dye-primer cycle sequencing protocol. The products of the sequencing reactions are run on sequencing gels and the sequences are determined using gel image analysis. The polymorphism search is based on the presence of superimposed peaks in the electrophoresis pattern resulting from different bases occurring at the same position. Because each dideoxy terminator is labeled with a different fluorescent molecule, the two peaks corresponding to a biallelic site present distinct colors corresponding to two different nucleotides at the same position on the sequence. However, the presence of two peaks can be an artifact due to background noise. To exclude such an artifact, the two DNA strands are sequenced and a comparison between the peaks is carried out. In order to be registered as a polymorphic sequence, the polymorphism has to be detected on both strands.

The above procedure permits those amplification products, which contain biallelic markers to be identified. The detection limit for the frequency of biallelic polymorphisms detected by sequencing pools of 100 individuals is approximately 0.1 for the minor allele, as verified by sequencing pools of known allelic frequencies. However, more than 90% of the biallelic polymorphisms detected by the pooling method have a frequency for the minor allele higher than 0.25. Therefore, the biallelic markers selected by this method have a frequency of at least 0.1 for the minor allele and less than 0.9 for the major allele. Preferably at least 0.2 for the minor allele and less than 0.8 for the major allele, more preferably at least 0.3 for the minor allele and less than 0.7 for the major allele, thus a heterozygosity rate higher than 0.18, preferably higher than 0.32, more preferably higher than 0.42.

In another embodiment, biallelic markers are detected by sequencing individual DNA samples, the frequency of the minor allele of such a biallelic marker may be less than 0.1.

Validation Of The Biallelic Markers Of The Present Invention

The polymorphisms are evaluated for their usefulness as genetic markers by validating that both alleles are present in a population. Validation of the biallelic markers is accomplished by genotyping a group of individuals by a method of the invention and demonstrating that both alleles are present. Microsequencing is a preferred method of genotyping alleles. The validation by genotyping step may be performed on individual samples derived from each individual in the group or by genotyping a pooled sample derived from more than one individual. The group can be as small as one individual if that individual is heterozygous for the allele in question. Preferably the group contains at least three individuals, more preferably the group contains five or six individuals, so that a single validation test will be more likely to result in the validation of more of the biallelic markers that are being tested. It should be noted,

These genotyping methods can be performed on nucleic acid samples derived from a single individual or pooled DNA samples.

Genotyping can be performed using similar methods as those described above for the identification of the biallelic markers, or using other genotyping methods such as those further described below. In preferred embodiments, the comparison of sequences of amplified genomic fragments from different individuals is used to identify new biallelic markers whereas microsequencing is used for genotyping known biallelic markers in diagnostic and association study applications.

In one embodiment the invention encompasses methods of genotyping comprising determining the identity of a nucleotide at a *PCTA-1*-related biallelic marker or the complement thereof in a biological sample; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A1 to A44, A46 to A53, A57, A58, A62 to A76, A81, A82, A86 to A91, A107, A118, and A123 to A125, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A45, A54, A60, A61, A77 to A80, A83 to A85, A93, A102 to A106, A109, A110, A114, and A122, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said *PCTA-1*-related biallelic marker is selected from the group consisting of A55, A56, A59, A92, A94 to A101, A108, A111 to A113, A115 to A117, and A119 to A121, and the complements thereof, or optionally the biallelic markers in linkage disequilibrium therewith; optionally, wherein said biological sample is derived from a single subject; optionally, wherein the identity of the nucleotides at said biallelic marker is determined for both copies of said biallelic marker present in said individual's genome; optionally, wherein said biological sample is derived from multiple subjects; Optionally, the genotyping methods of the invention encompass methods with any further limitation described in this disclosure, or those following, specified alone or in any combination; Optionally, said method is performed *in vitro*; optionally, further comprising amplifying a portion of said sequence comprising the biallelic marker prior to said determining step; Optionally, wherein said amplifying is performed by PCR, LCR, or replication of a recombinant vector comprising an origin of replication and said fragment in a host cell; optionally, wherein said determining is performed by a hybridization assay, a sequencing assay, a microsequencing assay, or an enzyme-based mismatch detection assay.

incorporated herein by reference in its entirety). As yet another alternative solid-phase microsequencing procedure, Nyren et al.(1993) described a method relying on the detection of DNA polymerase activity by an enzymatic luminometric inorganic pyrophosphate detection assay (ELIDA).

Pastinen et al.(1997) describe a method for multiplex detection of single nucleotide polymorphism in which the solid phase minisequencing principle is applied to an oligonucleotide array format. High-density arrays of DNA probes attached to a solid support (DNA chips) are further described below.

In one aspect the present invention provides polynucleotides and methods to genotype one or more biallelic markers of the present invention by performing a microsequencing assay. Preferred microsequencing primers include the nucleotide sequences D1 to D125 and E1 to E125. More preferred microsequencing primers are selected from the group consisting of the nucleotide sequences D15, D24, D30, D34, D36, D38, D41, D44, D50, D53, D54, D56, D57, D59, D76, D85, D93, D108, D111, D115, D124, E11, E14, E22, E25, E26, E35, E42, E52, E53, E55, E56, E60, E61, E64, E73, E75, E93, E96. It will be appreciated that the microsequencing primers listed in Example 4 are merely exemplary and that, any primer having a 3' end immediately adjacent to the polymorphic nucleotide may be used. Similarly, it will be appreciated that microsequencing analysis may be performed for any biallelic marker or any combination of biallelic markers of the present invention. One aspect of the present invention is a solid support which includes one or more microsequencing primers listed in Example 4, or fragments comprising at least 8, 12, 15, 20, 25, 30, 40, or 50 consecutive nucleotides thereof, to the extent that such lengths are consistent with the primer described, and having a 3' terminus immediately upstream of the corresponding biallelic marker, for determining the identity of a nucleotide at a biallelic marker site.

3) Mismatch detection assays based on polymerases and ligases

In one aspect the present invention provides polynucleotides and methods to determine the allele of one or more biallelic markers of the present invention in a biological sample, by mismatch detection assays based on polymerases and/or ligases. These assays are based on the specificity of polymerases and ligases. Polymerization reactions places particularly stringent requirements on correct base pairing of the 3' end of the amplification primer and the joining of two oligonucleotides hybridized to a target DNA sequence is quite sensitive to mismatches close to the ligation site, especially at the 3' end. Methods, primers and various parameters to amplify DNA fragments comprising biallelic markers of the present invention are further described above in "Amplification Of DNA Fragments Comprising Biallelic Markers".

herein by reference in its entirety, describes a tiling strategy for the detection of single nucleotide polymorphisms. Briefly, arrays may generally be "tiled" for a large number of specific polymorphisms. By "tiling" is generally meant the synthesis of a defined set of oligonucleotide probes which is made up of a sequence complementary to the target sequence of interest, as well as preselected variations of that sequence, e.g., substitution of one or more given positions with one or more members of the basis set of nucleotides. Tiling strategies are further described in PCT application No. WO 95/11995, the disclosure of which is incorporated herein by reference in its entirety. In a particular aspect, arrays are tiled for a number of specific, identified biallelic marker sequences. In particular, the array is tiled to include a number of detection blocks, each detection block being specific for a specific biallelic marker or a set of biallelic markers. For example, a detection block may be tiled to include a number of probes, which span the sequence segment that includes a specific polymorphism. To ensure probes that are complementary to each allele, the probes are synthesized in pairs differing at the biallelic marker. In addition to the probes differing at the polymorphic base, monosubstituted probes are also generally tiled within the detection block. These monosubstituted probes have bases at and up to a certain number of bases in either direction from the polymorphism, substituted with the remaining nucleotides (selected from A, T, G, C and U). Typically the probes in a tiled detection block will include substitutions of the sequence positions up to and including those that are 5 bases away from the biallelic marker. The monosubstituted probes provide internal controls for the tiled array, to distinguish actual hybridization from artefactual cross-hybridization. Upon completion of hybridization with the target sequence and washing of the array, the array is scanned to determine the position on the array to which the target sequence hybridizes. The hybridization data from the scanned array is then analyzed to identify which allele or alleles of the biallelic marker are present in the sample. Hybridization and scanning may be carried out as described in PCT application No. WO 92/10092 and WO 95/11995 and US patent No. 5,424,186, the disclosures of which are incorporated herein by reference in their entireties.

Thus, in some embodiments, the chips may comprise an array of nucleic acid sequences of fragments of about 15 nucleotides in length. In further embodiments, the chip may comprise an array including at least one of the sequences selected from the group consisting of amplicons listed in table 1 and the sequences complementary thereto, or a fragment thereof, said fragment comprising at least about 8 consecutive nucleotides, preferably 10, 15, 20, more preferably 25, 30, 40, 47, or 50 consecutive nucleotides and containing a polymorphic base. In preferred embodiments the polymorphic base is within 5, 4, 3, 2, 1, nucleotides of the center of the said polynucleotide, more preferably at the center of said polynucleotide. In some embodiments, the

used. Further, any set of genetic markers including a biallelic marker of the present invention may be used. A set of biallelic polymorphisms that could be used as genetic markers in combination with the biallelic markers of the present invention has been described in WO 98/20165, the disclosure of which is incorporated herein by reference in its entirety. As mentioned above, it should be noted that the biallelic markers of the present invention may be included in any complete or partial genetic map of the human genome. These different uses are specifically contemplated in the present invention and claims.

Linkage Analysis

Linkage analysis is based upon establishing a correlation between the transmission of genetic markers and that of a specific trait throughout generations within a family. Thus, the aim of linkage analysis is to detect marker loci that show cosegregation with a trait of interest in pedigrees.

Parametric Methods

When data are available from successive generations there is the opportunity to study the degree of linkage between pairs of loci. Estimates of the recombination fraction enable loci to be ordered and placed onto a genetic map. With loci that are genetic markers, a genetic map can be established, and then the strength of linkage between markers and traits can be calculated and used to indicate the relative positions of markers and genes affecting those traits (Weir, 1996). The classical method for linkage analysis is the logarithm of odds (lod) score method (see Morton, 1955; Ott, 1991). Calculation of lod scores requires specification of the mode of inheritance for the disease (parametric method). Generally, the length of the candidate region identified using linkage analysis is between 2 and 20Mb. Once a candidate region is identified as described above, analysis of recombinant individuals using additional markers allows further delineation of the candidate region. Linkage analysis studies have generally relied on the use of a maximum of 5,000 microsatellite markers, thus limiting the maximum theoretical attainable resolution of linkage analysis to about 600 kb on average.

Linkage analysis has been successfully applied to map simple genetic traits that show clear Mendelian inheritance patterns and which have a high penetrance (i.e., the ratio between the number of trait positive carriers of allele *a* and the total number of *a* carriers in the population). However, parametric linkage analysis suffers from a variety of drawbacks. First, it is limited by its reliance on the choice of a genetic model suitable for each studied trait. Furthermore, as already mentioned, the resolution attainable using linkage analysis is limited, and complementary studies are required to refine the analysis of the typical 2Mb to 20Mb regions initially identified through linkage analysis. In addition, parametric linkage analysis

Case-Control Populations (Inclusion Criteria)

Population-based association studies do not concern familial inheritance but compare the prevalence of a particular genetic marker, or a set of markers, in case-control populations. They are case-control studies based on comparison of unrelated case (affected or trait positive) individuals and unrelated control (unaffected, trait negative or random) individuals. Preferably the control group is composed of unaffected or trait negative individuals. Further, the control group is ethnically matched to the case population. Moreover, the control group is preferably matched to the case-population for the main known confusion factor for the trait under study (for example age-matched for an age-dependent trait). Ideally, individuals in the two samples are paired in such a way that they are expected to differ only in their disease status. The terms "trait positive population", "case population" and "affected population" are used interchangeably herein.

An important step in the dissection of complex traits using association studies is the choice of case-control populations (see Lander and Schork, 1994). A major step in the choice of case-control populations is the clinical definition of a given trait or phenotype. Any genetic trait may be analyzed by the association method proposed here by carefully selecting the individuals to be included in the trait positive and trait negative phenotypic groups. Four criteria are often useful: clinical phenotype, age at onset, family history and severity. The selection procedure for continuous or quantitative traits (such as blood pressure for example) involves selecting individuals at opposite ends of the phenotype distribution of the trait under study, so as to include in these trait positive and trait negative populations individuals with non-overlapping phenotypes. Preferably, case-control populations consist of phenotypically homogeneous populations. Trait positive and trait negative populations consist of phenotypically uniform populations of individuals representing each between 1 and 98%, preferably between 1 and 80%, more preferably between 1 and 50%, and more preferably between 1 and 30%, most preferably between 1 and 20% of the total population under study, and preferably selected among individuals exhibiting non-overlapping phenotypes. The clearer the difference between the two trait phenotypes, the greater the probability of detecting an association with biallelic markers. The selection of those drastically different but relatively uniform phenotypes enables efficient comparisons in association studies and the possible detection of marked differences at the genetic level, provided that the sample sizes of the populations under study are significant enough.

In preferred embodiments, a first group of between 50 and 300 trait positive individuals, preferably about 100 individuals, are recruited according to their phenotypes. A similar number of control individuals are included in such studies.

where genotype i occurs in phenotype j, and where h_k and h_l constitute genotype i. Each probability is derived according to eq. 1, and eq. 2 described above.

Then the Maximization step simply estimates another set of haplotype frequencies given the genotypes frequencies. This approach is also known as the gene-counting method (Smith, 1957).

$$p_t^{(s+1)} = \frac{1}{2} \sum_{j=1}^{F} \sum_{i=1}^{c_j} \delta_{it} . pr(genotype_i)^{(s)} \qquad \text{Equation 4}$$

Where δ_{it} is an indicator variable which count the number of time haplotype t in genotype i. It takes the values of 0, 1 or 2.

To ensure that the estimation finally obtained is the maximum-likelihood estimation several values of departures are required. The estimations obtained are compared and if they are different the estimations leading to the best likelihood are kept.

3) Methods To Calculate Linkage Disequilibrium Between Markers

A number of methods can be used to calculate linkage disequilibrium between any two genetic positions, in practice linkage disequilibrium is measured by applying a statistical association test to haplotype data taken from a population.

Linkage disequilibrium between any pair of biallelic markers comprising at least one of the biallelic markers of the present invention (M_i, M_j) having alleles (a_i/b_i) at marker M_i and alleles (a_j/b_j) at marker M_j can be calculated for every allele combination (a_i,a_j; a_i,b_j; b_i,a_j and b_i,b_j), according to the Piazza formula:

$\Delta_{aiaj} = \sqrt{}\theta4 - \sqrt{}(\theta4 + \theta3)(\theta4 + \theta2)$, where:

$\theta4$= - - = frequency of genotypes not having allele a_i at M_i and not having allele a_j at M_j

$\theta3$= - + = frequency of genotypes not having allele a_i at M_i and having allele a_j at M_j

$\theta2$= + - = frequency of genotypes having allele a_i at M_i and not having allele a_j at M_j

Linkage disequilibrium (LD) between pairs of biallelic markers (M_i, M_j) can also be calculated for every allele combination (ai,aj; ai,bj; b_i,a_j and b_i,b_j), according to the maximum-likelihood estimate (MLE) for delta (the composite genotypic disequilibrium coefficient), as described by Weir (Weir B. S., 1996). The MLE for the composite linkage disequilibrium is:

$D_{aiaj} = (2n_1 + n_2 + n_3 + n_4/2)/N - 2(pr(a_i).\ pr(a_j))$

Where $n_1 = \Sigma$ phenotype (a_i/a_i, a_j/a_j), $n_2 = \Sigma$ phenotype (a_i/a_i, a_j/b_j), $n_3 = \Sigma$ phenotype (a_i/b_i, a_j/a_j), n4= Σ phenotype (a_i/b_i, a_j/b_j) and N is the number of individuals in the sample.

individual. The mutations may lie within the coding sequence for the PCTA-1 protein or within intronic and/or within regulatory regions in the *PCTA-1* gene, including splice sites, 5' UTRs, 3' UTRs and promoter sequences, including one or more transcription factor binding sites.

A further embodiment of the invention is a method to identify a trait causing mutation in the *PCTA-1* gene pursuant to the detection of an association between alleles of one or several of the biallelic markers of the present invention and a particular trait.

This method comprises the following steps :

- amplifying a region of the *PCTA-1* gene comprising a biallelic marker or a group of biallelic markers associated to the considered trait from DNA samples of trait positive and trait negative individuals;
- sequencing the amplified region;
- comparing DNA sequences from trait positive and trait negative individuals; and
- determining mutations specific to trait positive patients.

In some embodiments, the amplified region is a region located close to a biallelic marker of *PCTA-1* gene. In further embodiments, the amplified region is located close to one or more of the biallelic markers A1 to A125 and the complements thereof. In a preferred embodiment, the amplified region is located close to one or more of the biallelic markers A2, A30, A41, A55, A57 and the complements thereof.

Oligonucleotide primers are constructed as described previously to amplify the sequences of each of the exons, introns, the promoter region and the regulatory regions of the *PCTA-1* gene. Amplification is carried out on genomic DNA samples from trait positive patients and trait negative controls, preferably using the PCR conditions described in the examples. Amplification products from the genomic PCRs are then subjected to sequencing, preferably through automated dideoxy terminator sequencing reactions and electrophoresed, preferably on ABI 377 sequencers. Following gel image analysis and DNA sequence extraction, ABI sequence data are automatically analyzed to detect the presence of sequence variations among trait positive and trait negative individuals. Sequences are verified by determining the sequences of both DNA strands for each individual.

Candidate polymorphisms suspected of being responsible for the detectable phenotype, are then verified by screening a larger population of trait positive and trait negative individuals using polymorphism analysis techniques such as the techniques described above. Polymorphisms which exhibit a statistically significant correlation with the detectable phenotype are deemed responsible for the detectable phenotype.

The invention also concerns a mutated *PCTA-1* gene comprising a trait causing mutation, and particularly the mutated genes obtained by the process described above.

production of the PCTA-1 protein, or with the production of a modified PCTA-1 protein. Diagnostic kits can comprise any of the polynucleotides of the present invention.

In a first embodiment, the kit comprises primers such as those described above, preferably forward and reverse primers which are used to amplify the *PCTA-1* gene or a fragment thereof. In some embodiments, at least one of the primers is complementary to a nucleotide sequence of the *PCTA-1* gene comprising a biallelic marker associated with prostate cancer, with an early onset of prostate cancer, with the level of aggressiveness of prostate cancer tumors, with a modified or forthcoming expression of the *PCTA-1* gene, with a modified or forthcoming production of the PCTA-1 protein, or with the production of a modified PCTA-1 protein. In one embodiment, the biallelic marker is comprised in one of the sequences P1 to P125 and the complementary sequences thereto. Optionally, the kit comprises an amplification primer which includes a polymorphic base of at least one biallelic marker selected from the group consisting of A1 to A125 and the complements thereof. In a preferred embodiment, the kit comprises one or more of the sequences B1 to B47 and C1 to C47. In a more preferred embodiment, the kit comprises one or more of the sequences B1, B16, B20, B23, B24 and C1, C16, C20, C23, C24.

In a second embodiment, the kit comprises microsequencing primers, wherein at least one of said primers is an oligonucleotide capable of hybridizing, either with the coding or with the non-coding strand, immediately upstream of the polymorphic base of a biallelic marker selected from the group consisting of A1 to A125 and the complements thereof, preferably those of D1 to D125 and E1 to E125, more preferably those of D2, D30, D41, D55, D57 and E2, E30, E41, E55, E57.

In a third embodiment, the kit comprises a hybridization DNA probe, that is or eventually becomes immobilized on a solid support, which is capable of hybridizing with the *PCTA-1* gene or fragment thereof, preferably which is capable of hybridizing with a region of the *PCTA-1* gene which comprises an allele of a biallelic marker of the present invention, more preferably an allele associated with prostate cancer, with an early onset of prostate cancer, with a susceptibility to prostate cancer, with the level of aggressiveness of prostate cancer tumors, with a modified or forthcoming expression of the *PCTA-1* gene, with a modified or forthcoming production of the PCTA-1 protein, or with the production of a modified PCTA-1 protein. In a preferred embodiment, the probe is selected from the group consisting of P1 to P125 and the complementary sequences thereto, or a fragment thereof, said fragment comprising the polymorphic base. In a more preferred embodiment, the probe is selected from the group consisting of P2, P30, P41, P55, P57 and the complementary sequences thereto, or a fragment thereof, said fragment comprising the polymorphic base.

The *in vivo* expression of a PCTA-1 polypeptide may be useful in order to correct a genetic defect related to the expression of the native gene in a host organism or to the production of a biologically inactive PCTA-1 protein.

Consequently, the present invention also deals with recombinant expression vectors mainly designed for the *in vivo* production of a PCTA-1 polypeptide of SEQ ID Nos 5, 6, 7, 9 or fragments or variants thereof by the introduction of the appropriate genetic material in the organism of the patient to be treated. This genetic material may be introduced *in vitro* in a cell that has been previously extracted from the organism, the modified cell being subsequently reintroduced in the said organism, directly *in vivo* into the appropriate tissue.

2. Regulatory Elements

Promoters

The suitable promoter regions used in the expression vectors according to the present invention are chosen taking into account the cell host in which the heterologous gene has to be expressed. The particular promoter employed to control the expression of a nucleic acid sequence of interest is not believed to be important, so long as it is capable of directing the expression of the nucleic acid in the targeted cell. Thus, where a human cell is targeted, it is preferable to position the nucleic acid coding region adjacent to and under the control of a promoter that is capable of being expressed in a human cell, such as, for example, a human or a viral promoter.

A suitable promoter may be heterologous with respect to the nucleic acid for which it controls the expression or alternatively can be endogenous to the native polynucleotide containing the coding sequence to be expressed. Additionally, the promoter is generally heterologous with respect to the recombinant vector sequences within which the construct promoter/coding sequence has been inserted.

Promoter regions can be selected from any desired gene using, for example, CAT (chloramphenicol transferase) vectors and more preferably pKK232-8 and pCM7 vectors.

Preferred bacterial promoters are the LacI, LacZ, the T3 or T7 bacteriophage RNA polymerase promoters, the gpt, lambda PR, PL and trp promoters (EP 0036776, the disclosure of which is incorporated herein by reference in its entirety), the polyhedrin promoter, or the p10 protein promoter from baculovirus (Kit Novagen) (Smith et al., 1983; O'Reilly et al., 1992), the lambda PR promoter or also the trc promoter.

Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionine-L. Selection of a convenient vector and promoter is well within the level of ordinary skill in the art.

Thus, the present invention also concerns a transgenic animal containing a nucleic acid, a recombinant expression vector or a recombinant host cell according to the invention.

Recombinant Cell Lines Derived From The Transgenic Animals Of The Invention.

A further object of the invention consists of recombinant host cells obtained from a transgenic animal described herein. In one embodiment the invention encompasses cells derived from non-human host mammals and animals comprising a recombinant vector of the invention or a *PCTA-1* gene disrupted by homologous recombination with a knock out vector.

Recombinant cell lines may be established *in vitro* from cells obtained from any tissue of a transgenic animal according to the invention, for example by transfection of primary cell cultures with vectors expressing *onc*-genes such as SV40 large T antigen, as described by Chou (1989) and Shay et al.(1991).

Screening Of Agents Acting Against Prostate Cancer

In a further embodiment, the present invention also concerns a method for the screening of new agents, or candidate substances, acting against cancer, preferably against prostate cancer and which may be suitable for the treatment of a patient whose DNA comprises an allele of the PCTA-1 gene associated with cancer, preferably with prostate cancer, with an early onset of prostate cancer, or with the aggressiveness of prostate cancer tumors, or more generally with a modified or forthcoming expression of the PCTA-1 gene, with a modified or forthcoming production of the PCTA-1 protein, or with the production of a modified PCTA-1 protein.

In a preferred embodiment, the invention relates to a method for the screening of candidate substances for cancer treatment, preferably prostate cancer treatment. The method comprises the following steps:

- providing a cell line, an organ, or a mammal expressing a *PCTA-1* gene or a fragment thereof, preferably the regulatory region or the promoter region of the *PCTA-1* gene;
- obtaining a candidate substance, preferably a candidate substance capable of inhibiting the binding of a transcription factor to the *PCTA-1* regulatory region; and
- testing the ability of the candidate substance to decrease the symptoms of cancer, preferably of prostate cancer and/or to modulate the expression levels of *PCTA-1*.

In some embodiments, the cell line, organ or mammal expresses a heterologous protein, the coding sequence of which is operably linked to the *PCTA-1* regulatory or promoter sequence. In other embodiments, they express a *PCTA-1* gene comprising alleles of one or more biallelic markers associated with cancer, preferably with prostate cancer, an early onset of prostate cancer, or the aggressiveness of prostate cancer tumors, or a mutated *PCTA-1* gene comprising a trait causing mutation determined using the above-noted method. Optionally, the

fragment thereof is fused to glutathion S transferase (GST). A mixture of cellular proteins or pool of expressed proteins as described above is applied to the affinity column. Proteins or other molecules interacting with the PCTA-1 protein or a fragment thereof attached to the column can then be isolated and analyzed on 2-D electrophoresis gel as described in Ramunsen et al. (1997). Alternatively, the proteins retained on the affinity column can be purified by electrophoresis based methods and sequenced. The same method can be used to isolate antibodies, to screen phage display products, or to screen phage display human antibodies.

D. Candidate Ligands Obtained By Optical Biosensor methods

Proteins interacting with the PCTA-1 protein or a fragment thereof can also be screened by using an Optical Biosensor as described in Edwards and Leatherbarrow (1997) and also in Szabo et al. (1995). This technique permits the detection of interactions between molecules in real time, without the need of labeled molecules. This technique is based on the surface plasmon resonance (SPR) phenomenon. Briefly, the candidate ligand molecule to be tested is attached to a surface (such as a carboxymethyl dextran matrix). A light beam is directed towards the side of the surface that does not contain the sample to be tested and is reflected by said surface. The SPR phenomenon causes a decrease in the intensity of the reflected light with a specific association of angle and wavelength. The binding of candidate ligand molecules cause a change in the refraction index on the surface, which change is detected as a change in the SPR signal. For screening of candidate ligand molecules or substances that are able to interact with the PCTA-1 protein or a fragment thereof, the PCTA-1 polypeptide is immobilized onto a surface. This surface consists of one side of a cell through which flows the candidate molecule to be assayed. The binding of the candidate molecule on the PCTA-1 protein or a fragment thereof is detected as a change of the SPR signal. The candidate molecules tested may be proteins, peptides, carbohydrates, lipids, or small molecules generated by combinatorial chemistry. This technique may also be performed by immobilizing eukaryotic or prokaryotic cells or lipid vesicles exhibiting an endogenous or a recombinantly expressed PCTA-1 protein at their surface.

The main advantage of the method is that it allows the determination of the association rate between the PCTA-1 protein and molecules interacting with the PCTA-1 protein. It is thus possible to select specifically ligand molecules interacting with the PCTA-1 protein, or a fragment thereof, through strong or conversely weak association constants.

E. Candidate Ligands Obtained Through A Two-Hybrid Screening Assay.

The yeast two-hybrid system is designed to study protein-protein interactions *in vivo* (Fields and Song, 1989), and relies upon the fusion of a bait protein to the DNA binding domain

of the yeast Gal4 protein. This technique is also described in the US Patent N° US 5,667,973 and the US Patent N° 5,283,173 (Fields et al.), the disclosures of which are incorporated herein by reference in their entireties.

The general procedure of library screening by the two-hybrid assay may be performed as described by Harper et al. (1993) or as described by Cho et al. (1998) or also Fromont-Racine et al. (1997).

The bait protein or polypeptide consists of a PCTA-1 polypeptide or a fragment thereof.

More precisely, the nucleotide sequence encoding the PCTA-1 polypeptide or a fragment thereof is fused to a polynucleotide encoding the DNA binding domain of the GAL4 protein, the fused nucleotide sequence being inserted in a suitable expression vector, for example pAS2 or pM3.

Then, a human cDNA library is constructed in a specially designed vector, such that the human cDNA insert is fused to a nucleotide sequence in the vector that encodes the transcriptional domain of the GAL4 protein. Preferably, the vector used is the pACT vector. The polypeptides encoded by the nucleotide inserts of the human cDNA library are termed "pray" polypeptides.

A third vector contains a detectable marker gene, such as beta galactosidase gene or CAT gene that is placed under the control of a regulation sequence that is responsive to the binding of a complete Gal4 protein containing both the transcriptional activation domain and the DNA binding domain. For example, the vector pG5EC may be used.

Two different yeast strains are also used. As an illustrative but non limiting example the two different yeast strains may be the followings :

Y190, the phenotype of which is (*MATa, Leu2-3, 112 ura3-12, trp1-901, his3-D200, ade2-101, gal4Dgal180D URA3 GAL-LacZ, LYS GAL-HIS3, cyh*^{r});

Y187, the phenotype of which is (*MATa gal4 gal80 his3 trp1-901 ade2-101 ura3-52 leu2-3, -112 URA3 GAL-lacZmet*$^-$), which is the opposite mating type of Y190.

Briefly, 20 µg of pAS2/*PCTA-1* and 20 µg of pACT-cDNA library are co-transformed into yeast strain Y190. The transformants are selected for growth on minimal media lacking histidine, leucine and tryptophan, but containing the histidine synthesis inhibitor 3-AT (50 mM). Positive colonies are screened for beta galactosidase by filter lift assay. The double positive colonies (*His*$^+$, *beta-gal*$^+$) are then grown on plates lacking histidine, leucine, but containing tryptophan and cycloheximide (10 mg/ml) to select for loss of pAS2/*PCTA-1* plasmids bu retention of pACT-cDNA library plasmids. The resulting Y190 strains are mated with Y187 strains expressing PCTA-1 or non-related control proteins; such as cyclophilin B, lamin, or SNF1, as *Gal4* fusions as described by Harper et al. (1993) and by Bram et al. (1993),

and screened for beta galactosidase by filter lift assay. Yeast clones that are *beta gal-* after mating with the control *Gal4* fusions are considered false positives.

In another embodiment of the two-hybrid method according to the invention, interaction between the PCTA-1 or a fragment thereof with cellular proteins may be assessed using the Matchmaker Two Hybrid System 2 (Catalog No. K1604-1, Clontech). As described in the manual accompanying the Matchmaker Two Hybrid System 2 (Catalog No. K1604-1, Clontech), nucleic acids encoding the PCTA-1 protein or a fragment thereof, are inserted into an expression vector such that they are in frame with DNA encoding the DNA binding domain of the yeast transcriptional activator GAL4. A desired cDNA, preferably human cDNA, is inserted into a second expression vector such that they are in frame with DNA encoding the activation domain of GAL4. The two expression plasmids are transformed into yeast and the yeast are plated on selection medium which selects for expression of selectable markers on each of the expression vectors as well as GAL4 dependent expression of the HIS3 gene. Transformants capable of growing on medium lacking histidine are screened for GAL4 dependent lacZ expression. Those cells which are positive in both the histidine selection and the lacZ assay contain interaction between PCTA-1 and the protein or peptide encoded by the initially selected cDNA insert.

Screening Through Spontaneous Metastatic Assay

Screening of new compounds can be carried out through a spontaneous metastatic assay as described in Nihei et al. (1995). Hence, it can be possible to assess the decrease of metastatic potential of transformed cells related to treatment of said compounds. Indeed, according to the present invention, the metastatic potential of cells constitutes the major criteria of the aggressiveness of prostate cancer tumors.

To evaluate the metastatic ability, about $5x10^5$ cells expressing a *PCTA-1* gene comprising alleles for one or more biallelic markers associated with cancer, preferably with prostate cancer, or with the aggressiveness of prostate cancer tumors, are injected subcutaneously in the flank of male athymic nude mice. The mice are treated with the screened compounds. Tumor volume and tumor volume doubling time are used as the index of the tumor growth rate and are determined as described in Isaacs & Hukku, 1988). The tumor-bearing animals are scored for lung metastases at spontaneous death or when killed at day 35 post-inoculation.

Gene Therapy

Gene therapy involves the alteration of the phenotypic expression of a targeted cell, usually a cancer cell through the alteration of the cell's genotypic content. The desired effect of gene therapy is a reduction or interruption of tumor growth or, ideally, the destruction of the cell

Vaccine composition

The invention concerns a vaccine composition comprising a vaccination agent including one of the following polypeptide:

a) a polypeptide comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids, more preferably at least 12, 15, 20, 25, 30, 40, 50, or 100 amino acids of SEQ ID No 5 , wherein said contigous span comprises:

i) a serine residue at amino acid position 170 and/or a lysine residue at amino acid position 203 in SEQ ID No 5; and/or

ii) at least one residue selected from the group consisting of a tyrosine residue at amino acid position 18, a cysteine residue at amino acid position 35, a methionine residue at amino acid position 55 and an arginine residue at amino acid position 183 in SEQ ID No 5;

b) a polypeptide comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids, more preferably at least 12, 15, 20, 25, 30, 40, 50, or 100 amino acids of SEQ ID No 6 , wherein said contigous span comprises:

i) a serine residue at amino acid position 170 and/or a lysine residue at amino acid position 245 in SEQ ID No 6; and/or

ii) at least one residue selected from the group consisting of a tyrosine residue at amino acid position 18, a cysteine residue at amino acid position 35, a methionine residue at amino acid position 55 and an arginine residue at amino acid position 225 in SEQ ID No 6; and/or

iii) at least 1, 2, 3, 5 or 10 of the amino acid encoded by the exon 6bis, more particularly at least 1, 2, 3, 5 or 10 of the amino acid positions 183-224 of the SEQ ID No 6;

c) a polypeptide comprising a contiguous span of at least 6 amino acids, preferably at least 8 to 10 amino acids, more preferably at least 12, 15, 20, 25, 30, 40, 50, or 100 amino acids of SEQ ID No 7 , wherein said contigous span comprises:

i) a serine residue at amino acid position 170 and/or a lysine residue at amino acid position 203 in SEQ ID No 7; and/or

ii) at least one residue selected from the group consisting of a tyrosine residue at amino acid position 18, a cysteine residue at amino acid position 35, a methionine residue at amino acid position 55 and an arginine residue at amino acid position 183 in SEQ ID No 7; and/or

Various methods of achieving adjuvant effects for vaccines include the use of agents such as aluminim hydroxide or phosphate, commonly used as 0.05 to 0.1 percent solution in phosphate buffered saline, admixture with synthetic polymers of sugars (Carbopol) used as 0.25 percent solution, aggregation of the protein in the vaccine by heat treatment with temperature ranging between 70°C and 101°C for 30 second to 2 minute periods, respectively. Aggregation by reacting with pepsin treated antibodies (Fab) to albumin, mixture with bacterial cells such as C. parvum or endotoxins or lipopolysaccharide components of gram-negative bacteria, emulsion in physiologically acceptable oil vehicles such as mannide monoleate (Aracel A) or emulsion with 20 percent solution of a perfluorocarbon (Fluosol-DA) used as a block substitute may also be employed. According to the invention, dimethyldioctadecylammonium bromide is an interesting candidate for an adjuvant, but also Freund's complete and incomplete adjuvants as well as QuilA and RIBI are interesting possibilities. Other possibilities involve the use of immune modulating substances such as lymphokines (e.g. IFN-γ, IL-2 and IL-12) or synthetic IFN-γ inducers such as poly I:C in combination with the above-mentioned adjuvants.

The vaccine agent of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby immunogenic peptides, or their functional derivatives, are combined in admixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation, inclusive of other human proteins, such as human serum albumin, are described Remington's Pharmaceutical Sciences (1980). In order to form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of one or more of the vaccine agents of the present invention, together with a suitable amount of a carrier vehicle.

Additional pharmaceutical methods may be employed to control the duration of action. Control release preparations may be achieved through the use of polymers to complex or absorb one or more of the vaccine agents of the present invention. The controlled delivery may be exercised by selecting appropriate macromolecule (for example polyesters, polyamino acids, polyvinyl, pyrrolidone, ethylenevinylacetate, methylcellulose, carboxymethylcellulose, protamine, or sulfate) and the concentration of macromolecules as well as the methods of incorporation in order to control release. Another possible method to control the duration of action by controlled release preparations is to incorporate vaccine agents of the present invention into particles of a polymeric material such as polyesters, polyamino acids, hydrogels, poly(lactic acid) or ethylene vinylacetate copolymers. Alternatively, instead of incorporating these vaccine agents into polymeric particles, it is possible to entrap these materials in microcapsules prepared, for example, bu coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatine-microcapsules and

media may be a hard disk, a floppy disk, a magnetic tape, CD-ROM, Digital Versatile Disk (DVD), Random Access Memory (RAM), or Read Only Memory (ROM) as well as other types of other media known to those skilled in the art.

Embodiments of the present invention include systems, particularly computer systems which store and manipulate the sequence information described herein. One example of a computer system 100 is illustrated in block diagram form in Figure 3. As used herein, "a computer system" refers to the hardware components, software components, and data storage components used to analyze the nucleotide sequences of the nucleic acid codes of the invention or the amino acid sequences of the polypeptide codes of the invention. In one embodiment, the computer system 100 is a Sun Enterprise 1000 server (Sun Microsystems, Palo Alto, CA). The computer system 100 preferably includes a processor for processing, accessing and manipulating the sequence data. The processor 105 can be any well-known type of central processing unit, such as the Pentium III from Intel Corporation, or similar processor from Sun, Motorola, Compaq or International Business Machines.

Preferably, the computer system 100 is a general purpose system that comprises the processor 105 and one or more internal data storage components 110 for storing data, and one or more data retrieving devices for retrieving the data stored on the data storage components. A skilled artisan can readily appreciate that any one of the currently available computer systems are suitable.

In one particular embodiment, the computer system 100 includes a processor 105 connected to a bus which is connected to a main memory 115 (preferably implemented as RAM) and one or more internal data storage devices 110, such as a hard drive and/or other computer readable media having data recorded thereon. In some embodiments, the computer system 100 further includes one or more data retrieving device 118 for reading the data stored on the internal data storage devices 110.

The data retrieving device 118 may represent, for example, a floppy disk drive, a compact disk drive, a magnetic tape drive, etc. In some embodiments, the internal data storage device 110 is a removable computer readable medium such as a floppy disk, a compact disk, a magnetic tape, etc. containing control logic and/or data recorded thereon. The computer system 100 may advantageously include or be programmed by appropriate software for reading the control logic and/or the data from the data storage component once inserted in the data retrieving device.

The computer system 100 includes a display 120 which is used to display output to a computer user. It should also be noted that the computer system 100 can be linked to other computer systems 125a-c in a network or wide area network to provide centralized access to the computer system 100.

Once a comparison of the two sequences has been performed at the state 210, a determination is made at a decision state 210 whether the two sequences are the same. Of course, the term "same" is not limited to sequences that are absolutely identical. Sequences that are within the homology parameters entered by the user will be marked as "same" in the process 200.

If a determination is made that the two sequences are the same, the process 200 moves to a state 214 wherein the name of the sequence from the database is displayed to the user. This state notifies the user that the sequence with the displayed name fulfills the homology constraints that were entered. Once the name of the stored sequence is displayed to the user, the process 200 moves to a decision state 218 wherein a determination is made whether more sequences exist in the database. If no more sequences exist in the database, then the process 200 terminates at an end state 220. However, if more sequences do exist in the database, then the process 200 moves to a state 224 wherein a pointer is moved to the next sequence in the database so that it can be compared to the new sequence. In this manner, the new sequence is aligned and compared with every sequence in the database.

It should be noted that if a determination had been made at the decision state 212 that the sequences were not homologous, then the process 200 would move immediately to the decision state 218 in order to determine if any other sequences were available in the database for comparison.

Accordingly, one aspect of the present invention is a computer system comprising a processor, a data storage device having stored thereon a nucleic acid code of the invention or a polypeptide code of the invention, a data storage device having retrievably stored thereon reference nucleotide sequences or polypeptide sequences to be compared to the nucleic acid code of the invention or polypeptide code of the invention and a sequence comparer for conducting the comparison. The sequence comparer may indicate a homology level between the sequences compared or identify structural motifs in the nucleic acid code of the invention and polypeptide codes of the invention or it may identify structural motifs in sequences which are compared to these nucleic acid codes and polypeptide codes. In some embodiments, the data storage device may have stored thereon the sequences of at least 2, 5, 10, 15, 20, 25, 30, or 50 of the nucleic acid codes of the invention or polypeptide codes of the invention.

Another aspect of the present invention is a method for determining the level of homology between a nucleic acid code of the invention and a reference nucleotide sequence, comprising the steps of reading the nucleic acid code and the reference nucleotide sequence through the use of a computer program which determines homology levels and determining homology between the nucleic acid code and the reference nucleotide sequence with the computer program. The computer program may be any of a number of computer programs for determining homology levels,

Example 2

Detection Of The Biallelic Markers: Amplification Of Genomic DNA By PCR

The amplification of specific genomic sequences of the DNA samples of example 1 was carried out on the pool of DNA obtained previously. In addition, 10 individual samples were similarly amplified.

PCR assays were performed using the following protocol:

Final volume	25 µl
DNA	2 ng/µl
$MgCl_2$	2 mM
dNTP (each)	200 µM
primer (each)	2.9 ng/µl
Ampli Taq Gold DNA polymerase	0.05 unit/µl
PCR buffer (10x = 0.1 M TrisHCl pH8.3 0.5M KCl	1x

Each pair of primers was designed using the sequence information of our total genomic sequence (SEQ ID No 1) and the OSP software (Hillier & Green, 1991). These primers had about 20 nucleotides in length and their respective sequences are disclosed in Table 1 and had the sequences disclosed in Table 1 in the columns labeled "Position range of amplification primer in SEQ ID No 1" and "Complementary position range of amplification primer in SEQ ID No 1".

The primers contained a common oligonucleotide tail upstream of the specific bases targeted for amplification which was useful for sequencing.

Primers from the columns labeled "Position range of amplification primer in SEQ ID No 1," contain the following additional PU 5' sequence: TGTAAAACGACGGCCAGT; and primers from the columns labeled "Complementary position range of amplification primer in SEQ ID No 1," contain the following RP 5' sequence: CAGGAAACAGCTATGACC. The primer containing the additional PU 5' sequence is listed in SEQ ID No 10. The primer containing the additional RP 5' sequence is listed in SEQ ID No 11.

The synthesis of these primers was performed following the phosphoramidite method, on a GENSET UFPS 24.1 synthesizer.

DNA amplification was performed on a Genius II thermocycler. After heating at 94°C for 10 min, 40 cycles were performed. Each cycle comprised: 30 sec at 94°C, 55°C for 1 min, and 30 sec at 72°C. For final elongation, 7 min at 72°C end the amplification. The quantities of the amplification products obtained were determined on 96-well microtiter plates, using a fluorometer and Picogreen as intercalant agent (Molecular Probes).

Table 6: Haplotype frequency analysis for the sporadic cases of prostate cancer

			A2	A30	A41	A55	A57	A75	haplotype frequencies		Odds ratio	Pvalue (1df)
frequency%			60/67 (A)	64/66 (T)	73/71 (T)	64/68 (C)	65/69 (G)	94/95 (G)	cases	controls		
abs diff freq. all			-7,4	-2,0	2,7	-3,7	-3,9	-1				
pvalue			7,7e-03	4,3e-01	2,9e-01	1,6e-01	1,4e-01	3,4e-01				
Cases/controls ↓												
H1	PT2	281/298	T	1		T			**0.162**	**0.090**	**1.94**	**2.4e-04**
H2		282/301	T	2			A		**0.163**	**0.095**	**1.85**	**5.3e-04**
H3		282/301		3	T	T			0.140	0.083	1.79	2.1e-03
H4		283/298		4	T		A		0.136	0.083	1.74	3.6e-03
H6		283/299	T	5	T				0.317	0.246	1.42	7.3e-03
H7		279/300		6 T				T	0.045	0.022	2.08	3.0e-02
H1	PT3	278/295	T	7	T	T			0.083	0.037	2.33	1.1e-03
H2		277/294	T	8		T	A		**0.150**	**0.092**	**1.75**	**2.3e-03**
H3		279/295	T	9	T		A		0.081	0.040	2.12	3.4e-03
H4		278/294		10	T	T	A		0.134	0.082	1.75	3.8e-03
H6		277/295	T	11		T		G	0.126	0.076	1.76	4.7e-03
H7		277/293		12 T	T		A		0.091	0.048	1.96	4.7e-03
H8		275/294		13 T	T	T			0.093	0.051	1.91	5.5e-03
H1	PT4	273/290	T	14 T	T		A		0.046	0.010	4.76	2.0e-04
H2		271/290	T	15 T	T	T			0.044	0.010	4.54	3.9e-04
H3		274/291	T	16	T	T	A		0.078	0.038	2.15	3.6e-03
H4		274/292	T	17	T	T		G	0.053	0.021	2.57	4.4e-03
H6		272/289		18 T	T	T	A		0.090	0.048	1.95	5.5e-03

Table 7: Haplotype frequency analysis of the preferred haplotypes

		HAPLOTYPE						Pvalue haplo. Frequency % (cases vs controls)	
	MARKERS		A2	A30	A41	A55	A57	familial cases vs controls	sporadic cases vs controls
FAMILIAL CASES HAPLOTYPES	**PT2**	**haplotype 1**		**T**	**T**			**1e-04 (57/44)**	6e-01 (45/44)
	PT3	**haplotype 2**	**A**	**T**	**T**			**1e-05 (43/29)**	2e-01 (26/29)
SPORADICS CASES HAPLOTYPES	**PT2**	**haplotype 3**	**T**				**A**	4e-01 (11/10)	**5e-04 (16/10)**
		haplotype 4	**T**			**T**		3e-01 (11/9)	**2e-04 (16/9)**
	PT3	**haplotype 5**	**T**			**T**	**A**	3e-01 (11/9)	**2e-03 (15/9)**

5

ELISA, as originally described by Engvall, (1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis et al.

B. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes in a PCTA-protein or a portion thereof can be prepared by immunizing suitable non-human animal with this PCTA-1 protein or a portion thereof, which can be unmodified or modified to enhance immunogenicity. A suitable non-human animal is preferably a non-human mammal is selected, usually a mouse, rat, rabbit, goat, or horse. Alternatively, a crude preparation which has been enriched for the PCTA-1 concentration can be used to generate antibodies. Such proteins, fragments or preparations are introduced into the non-human mammal in the presence of an appropriate adjuvant (e.g. aluminum hydroxide, RIBI, etc.) which is known in the art. In addition the protein, fragment or preparation can be pretreated with an agent which will increase antigenicity, such agents are known in the art and include, for example, methylated bovine serum albumin (mBSA), bovine serum albumin (BSA), Hepatitis B surface antigen, and keyhole limpet hemocyanin (KLH). Serum from the immunized animal is collected, treated and tested according to known procedures. If the serum contains polyclonal antibodies to undesired epitopes, the polyclonal antibodies can be purified by immunoaffinity chromatography.

Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. Techniques for producing and processing polyclonal antisera are known in the art, see for example, Mayer and Walker (1987). An effective immunization protocol for rabbits can be found in Vaitukaitis, et al. (1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. et al., (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, (1980).

Antibody preparations prepared according to either the monoclonal or the polyclonal protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively

Livak et al., *Nature Genetics*, 9:341-342, 1995

Livak KJ, Hainer JW, *Hum Mutat* 1994;**3**(4):379-385

Lockhart et al. *Nature Biotechnology* **14**: 1675-1680, 1996

Lucas A.H., 1994, In : Development and Clinical Uses of Haempophilus b Conjugate;

Manley JL, et al, *Proc Natl Acad Sci U S A.* 1980 Jul; **77**(7): 3855-3859

Mansour S.L. et al. (1988) *Nature.* 336:348-352.

Marshall R. L. et al. (1994) *PCR Methods and Applications.* 4:80-84.

Maxam AM, Gilbert W, *Methods Enzymol* 1980;**65**:499-560

McCormick et al. (1994) *Genet. Anal. Tech. Appl.* 11:158-164.

McLaughlin B.A. et al. (1996) *Am. J. Hum. Genet.* 59:561-569.

Mizokami A, Yeh SY, Chang C, *Mol Endocrinol* 1994;**8**:77-88

Morton N.E., *Am.J. Hum.Genet.*, 7:277-318, 1955

Muller MM, Schreiber E, Schaffner W, Matthias P, *Nucleic Acids Res* 1989;**17**:6420

Muzyczka et al. (1992) *Curr. Topics in Micro. and Immunol.* 158:97-129.

Nada S. et al. (1993) *Cell* 73:1125-1135.

Nagy A. et al., 1993, Proc. Natl. Acad. Sci. USA, **90**: 8424-8428.

Narang SA, Hsiung HM, Brousseau R, *Methods Enzymol* 1979;**68**:90-98

Neda et al. (1991) *J. Biol. Chem.* 266:14143-14146.

Newton et al. (1989) *Nucleic Acids Res.* 17:2503-2516.

Nickerson D.A. et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:8923-8927.

Nicolau C. et al., 1987, Methods Enzymol., **149**:157-76.

Nicolau et al. (1982) *Biochim. Biophys. Acta.* 721:185-190.

Nihei et al, *Genes Chromosomes Cancer* 1995;**14**:112-119

Nyren P, Pettersson B, Uhlen M, *Anal Biochem* 1993;**208**(1):171-175

O'Reilly et al. (1992) *Baculovirus Expression Vectors: A Laboratory Manual.* W. H. Freeman and Co., New York.

Ohno et al. (1994) *Science.* 265:781-784.

Oldenburg K.R. et al., 1992, Proc. Natl. Acad. Sci., **89**:5393-5397.

Orita et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.*86: 2776-2770.

Ott J., *Analysis of Human Genetic Linkage, John Hopkins University Press, Baltimore,* 1991

Ouchterlony, O. et al., Chap. 19 in: Handbook of Experimental Immunology D. Wier (ed) Blackwell (1973)

Parmley and Smith, Gene, 1988, **73**:305-318

Pastinen et al., *Genome Research* 1997; 7:606-614

Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85(8):2444-2448

Shizuya et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:8794-8797.

Shoemaker DD, et al., *Nat Genet* 1996;**14**(4):450-456

Siebenlist U, Gilbert W, *Proc Natl Acad Sci U S A* 1980;**77**:122-126

Smith (1957) *Ann. Hum. Genet.* 21:254-276.

Smith et al. (1983) *Mol. Cell. Biol.* 3:2156-2165.

Sosnowski RG, et al., *Proc Natl Acad Sci U S A* 1997;**94**:1119-1123

Sowdhamini et al., Protein Engineering 10:207, 215 (1997)

Spielmann S. and Ewens W.J., *Am. J. Hum. Genet.*, 62:450-458, 1998

Spielmann S. et al., *Am. J. Hum. Genet.*, 52:506-516, 1993

Sternberg N.L. (1994) *Mamm. Genome.* 5:397-404.

Strittmatter WJ, et al., *Proc Natl Acad Sci U S A* 1993 ;**90**:1977-1981

Stryer, L., *Biochemistry,* 4th edition, 1995

Syvanen AC, *Clin Chim Acta* 1994;**226**(2):225-236

Szabo A. et al. *Curr Opin Struct Biol* **5**, 699-705 (1995)

Szabo et al., 1995, Curr Opin Struct Biol., **5**(5):699-705

Tacson et al. (1996) *Nature Medicine.* 2(8):888-892.

Te Riele et al. (1990) Nature. 348:649-651.

Terwilliger J.D. and Ott J., *Handbook of Human Genetic Linkage, John Hopkins University Press, London*, 1994

Thomas K.R. et al. (1986) *Cell.* 44:419-428.

Thomas K.R. et al. (1987) *Cell.* 51:503-512.

Thompson et al., 1994, Nucleic Acids Res. 22(2):4673-4680

Tur-Kaspa et al. (1986) *Mol. Cell. Biol.* 6:716-718.

Tyagi et al. (1998) *Nature Biotechnology.* 16:49-53.

Urdea M.S. (1988) *Nucleic Acids Research.* 11:4937-4957.

Urdea M.S. et al.(1991) *Nucleic Acids Symp. Ser.* 24:197-200.

Vaitukaitis, J. et al. J. Clin. Endocrinol. Metab. 33:988-991 (1971)

Valadon P., et al., 1996, J. Mol. Biol., **261**:11-22.

Van der Lugt et al. (1991) *Gene.* 105:263-267.

Vlasak R. et al. (1983) *Eur. J. Biochem.* 135:123-126.

Wabiko et al. (1986) *DNA.*5(4):305-314.

Walker et al. (1996) *Clin. Chem.* 42:9-13.

Wang et al., 1997, Chromatographia, **44** : 205-208.

Weir, B.S. (1996) *Genetic data Analysis II: Methods for Discrete population genetic Data, Sinauer Assoc., Inc., Sunderland, MA, U.S.A.*

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450 Alexandria, VA 22313 on January 14, 2005.

Frank C. Eisenschenk

Frank C. Eisenschenk, Ph.D., Patent Attorney

REQUEST FOR CERTIFICATE OF CORRECTION UNDER 37 CFR 1.322 and 1.323
Docket No. GEN-T112XC1
Patent No. 6,759,192

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov
Issued	:	July 6, 2004
Patent No.	:	6,759,192
For	:	Polymorphic Markers of the Prostate Carcinoma Tumor Antigen-1 (PCTA-1)

ATTN: CERTIFICATE OF CORRECTIONS BRANCH
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR CERTIFICATE OF CORRECTION
UNDER 37 CFR 1.322 (OFFICE MISTAKE) AND
UNDER 37 CFR 1.323 (APPLICANTS' MISTAKE)

Sir:

A Certificate of Correction for the above-identified patent has been prepared and is attached hereto.

In the left-hand column below is the column and line number where errors occurred in the patent. In the right-hand column is the page and line number in the application where the correct information appears.

COPY

Patent Reads:	**Application Reads:**
Column 4, Line 24: "and (consenus"	See Amendment dated 12/15/03, page 2: --and 22 (consensus)--
Column 4, Line 59: "82393. SEQ ID NO: 12"	See Examiner's Amendment dated 02/24/04, page 3: --82393.--
Column 5, Line 60: "Polynucleotides"	Page 7, Line 2: --polynucleotides--
Column 7, Line 17: "Volypeptide"	Page 8, Line 26: --polypeptide--
Column 8, Line 27: "heterozyposity"	Page 10, Line 6: --heterozygosity rate--
Column 11, Line 14: "bronchial En asthma"	Page 13, Lines 20-21: --bronchial asthma--
Column 14, Line 1: "BLASr"	Page 17, Line 8: --BLAST--
Column 15, Line 15: "exon 7, exon 9"	Page 18, Lines 29-30: --exon 7, exon 8, exon 9--
Column 20, Line 7: "nucleotide c at"	Page 24, Lines 30-31: --nucleotide C at--
Column 25, Line 8: "4400-8000"	Page 30, Line 24: --44001-48000--
Column 25, Line 44: "3'-regulatory,polynucleotide"	Page 31, Line 12: --3'-regulatory polynucleotide --

Column 27, Line 18: "seven let"	Page 33, Line 17: --seven *tet*--
Column 27, Line 34: "the let operator"	Page 33, Line 27: --the *tet* operator--
Column 31, Line 12: "24042406"	Page 38, Line 13: --2404-2406--
Column 31, Line 49: "nucleotide Tat positions"	Page 38, Line 35: --nucleotide T at positions--
Column 32, line 62: "thereto;"	Page 40, Lines 16-17: --thereto; and--
Column 33, Line 55: "A118, A123"	Page 41, Lines 20-21: --A118, and A123--
Column 34, Line 42: "A92A94to"	Page 42, Line 20: --A92, A94 to --
Column 35, Line 3: "A1107"	Page 43, Line 3: --A107--
Column 35, Line 9: "and complements"	Page 43, Lines 6-7: --and the complements--
Column 35, Line 13: "A55A56A59A92A94to"	Page 43, Line 9: --A55, A56, A59, A92, A94 to--
Column 35, Line 13: "A108A111 to A113A115"	Page 43, Line 9: --A108, A111 to A113, A115--
Column 36, Line 13: "^{32}p"	Page 44, Line 19: --^{32}P--
Column 38, Line 45: "VLSIPS*v*"	Page 47, Line 25: --VLSIPS™--

Column 50, Line 16:
"A123 complements"

Column 50, Line 20:
"A114, A122"

Column 50, Line 25:
"to 113, A115, A117, and"

Column 50, Line 64:
"A123 complements"

Column 51, Line 1:
"A109, A122,"

Column 52, Line 10:
"are Q selected"

Column 56, Line 48:
"the w biallelic"

Column 58, Line 25:
"A109, A122"

Column 58, Line 29:
"A94 to A01,"

Column 62, Line 9:
"Eli"

Column 66, Line 25:
"about IS"

Column 67, Line 27:
"WO 98120165"

Column 73, Lines 22-23:
"with overlapping"

Page 62, Lines 10-11:
--A123 to A125, and the complements"--

Page 62, Lines 13-14:
--A114, and A122--

Page 62, Lines 16-17:
--to A113, A115 to A117, and--

Page 63, Lines 6-7:
--A123 to A125, and the complements--

Page 63, Lines 9-10:
--A109, A110, A114, and A122--

Page 64, Lines 22-23:
--are selected--

Page 70, Lines 15-16:
--the biallelic--

Page 72, Lines 18-19:
--A109, A110, A114, and A122--

Page 72, Line 21:
--A94 to A101,--

Page 77, Line 14:
--E11--

Page 82, Line 29:
--about 15--

Page 84, Line 4:
--WO 98/20165--

Page 91, Lines 21-22:
--with non-overlapping--

Column 87, Line 2:
"(a_i,a_j, a_i,b_j: b_i,a_j)"

Page 108, Line 19:
--(a_i,a_j; a_i,b_j; b_i,a_j)--

Column 87, Line 7:
"θ4=—=frequency"

Page 108, Line 22:
--θ4= - - = frequency--

Column 87, Line 9:
"θ3=—+=frequency"

Page 108, Line 23:
--θ3= - + = frequency--

Column 87, Line 11:
"θ2=+—=frequency"

Page 108, Line 24:
--θ2= + - = frequency--

Column 87, Line 14:
"(M_i; M_j)"

Page 108, Line 26:
--(M_i, M_j)--

Column 87, Line 21:
"D_{aiaj} ($2n_1$+"

Page 108, Line 30:
--D_{aiaj}= ($2n_1$+--

Column 92, Line 29:
"AS5"

Page 115, Line 18:
--A55--

Column 95, Lines 41-42:
"preferred 1 5 embodiment"

Page 119, Lines 14-15:
--preferred embodiment--

Column 101, Line 49:
"be I,X;1 expressed"

Page 127, Lines 13-14:
--be expressed--

Column 109, Line 22:
"one-genes"

Page 137, Line 10:
--*onc*-genes--

Column 117, Line 15:
"(ISPR)"

Page 147, Line 13:
--(SPR)--

Column 117, Line 17:
"carboxymethyi"

Page 147, Line 14:
--carboxymethyl--

Column 118, Line 16:
"(MaTa gal4"

Page 148, Line 25:
--(MATa gal4--

Column 118, Line 18: "lacZmet'),""	Page 148, Line 26: --lacZmet-)--
Column 118, Line 52: "GALA"	Line 149, Line 13: --GAL4--
Column 125, Line 6: "more CF particularly"	Page 157, Lines 22-23: --more particularly--
Column 127, Line 57: "monoleatc"	Page 161, Line 8: --monoleate--
Column 132, Line 42: "CAROM"	Page 167, Line 1: --CD-ROM--
Column 134, Line 47: "or a By polypeptide"	Page 169, Lines 20-21: --or a polypeptide--
Column 136, Line 47: "TTATAA Box"	See Amendment dated 02/24/03, page 2: --TAATAA Box--
Column 136, Line 48: "TATAA"	See Amendment dated 02/24/03, page 2: --TAATAA--
Column 140, Line 10: "[BLANK]"	Page 176, Line 6: --PCR assays were performed using the following protocol:--
Column 158, Line 1 of Table 6: "analysis of the"	Page 195, Line 1: --analysis for the--

Column 159, Line 2:
"analysis of the"

Page 195, Line 1:
--analysis for the--

Column 159, Line 6 of Table 6:
"

"

Page 195, Lines 3-5:
--

	A2	A30	A41	A55	A57	A75
frequency %	60/67 (A)	64/66 (T)	73/71 (T)	64/68 (C)	65/69 (G)	94/95 (G)
abs diff freq, all	-7,4	-2,0	2,7	-3,7	-3,9	-1
pvalue	7,7e-03	4,3e-01	2,9e-01	1,6e-01	1,4e-01	3,4e-01
Cases/controls ↓						

--

Column 164, Line 14:
"and b dose"

Page 199, Lines 20-21:
--and dose--

Column 164, Line 34:
"they arc"

Page 199, Line 35:
--they are--

Column 167, Line 25:
"Baltimore, 1991 Ouchterlony, O."

Page 204, Lines 30-31:
--Baltimore, 1991
Ouchterlony, O.--

Column 168, Line 22:
"119-1123"

Page 206, Line 6:
--1119-1123--

Column 168, Line 26:
"Stembert"

Page 206, Line 10:
--Sternberg--

Column 168, Line 26:
"5:397404"

Page 206, Line 10:
--5:397-404--

Column 168, Line 27:
"Strinmatter"

Patent Reads:

Column 569, Line 28:
"and b) performed"

Column 570, Line 25:
"A75 and A30;"

Page 206, Line 11:
--Strittmatter--

Application Should Read:

See Amendment dated 12/15/03, page 6, line 6, Claim 137:
--and b) are performed--

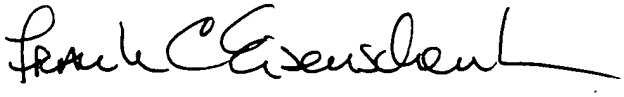
See Amendment dated 12/15/03, page 7, line 11, Claim 150:
--A75;--

True and correct copies of pages 7, 8, 10, 13, 17, 18, 24, 30, 31, 33, 37, 38, 40, 41, 42, 43, 44, 47, 62, 63, 64, 70, 72, 77, 82, 84, 91, 108, 115, 119, 127, 137, 147, 148, 149, 157, 161, 167, 169, 176, 195, 199, 204, and 206 of the specification as filed; Applicant's Amendments dated February 24, 2003 and December 15, 2003; and the Examiner's Amendment dated February 24, 2004 which support Applicants' assertion of errors on the part of the Patent Office accompanies this Certificate of Correction.

The Commissioner is authorized to charge the fee of $100.00 for the amendment to Deposit Account No. 19-0065. The Commissioner is also authorized to charge any additional fees as required under 37 CFR 1.20(a) to Deposit Account No. 19-0065. Two copies of this letter are enclosed for Deposit Account authorization.

Approval of the Certificate of Correction is respectfully requested.

Respectfully submitted,



Frank C. Eisenschenk, Ph.D.
Patent Attorney
Registration No. 45,332
Phone No.: 352-375-8100
Fax No.: 352-372-5800
Address: P.O. Box 142950
Gainesville, FL 32614-2950

FCE/mv
Attachments: Certificate of Correction
Copies of pages 7, 8, 10, 13, 17, 18, 24, 30, 31, 33, 37, 38, 40, 41, 42, 43, 44, 47, 62, 63, 64, 70, 72, 77, 82, 84, 91, 108, 115, 119, 127, 137, 147, 148, 149, 157, 161, 167, 169, 176, 195, 199, 204, and 206 of the specification as filed; the applicant's Amendments dated February 24, 2003 and December 15, 2003; and the Examiner's Amendment dated February 24, 2004
Two copies of this letter

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 4
Line 24, "and (consenus" should read --and 22 (consensus)--.

Line 59, "82393. SEQ ID NO: 12" should read --82393.--.

Column 5
Line 60, "Polynucleotides" should read --polynucleotides--.

Column 7
Line 17, "Volypeptide" should read --polypeptide--.

Column 8
Line 27, "heterozyposity" should read --heterozygosity rate--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

FORM PTO-1050 (REV. 3-75) UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 11
Line 14, "bronchial En asthma" should read --bronchial asthma--.

Column 14
Line 1, "BLASr" should read --BLAST--.

Column 15
Line 15, "exon 7, exon 9" should read --exon 7, exon 8, exon 9--.

Column 20
Line 7, "nucleotide c at" should read --nucleotide C at--.

Column 25
Line 8, "4400-8000" should read --44001-48000--.

Line 44, "3'-regulatory,polynucleotide" should read --3'-regulatory polynucleotide --.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies

@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 27
Line 18, "seven let" should read --seven *tet*--.

Line 34, "the let operator" should read --the *tet* operator--.

Column 31
Line 12, "24042406" should read --2404-2406--.

Line 49, "nucleotide Tat positions" should read --nucleotide T at positions--.

Column 32
Line 62, "thereto;" should read --thereto; and--.

Column 33
Line 55, "A118, A123" should read --A118, and A123--

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192
No. of add'l. copies
@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 34
Line 42, "A92A94to" should read --A92, A94 to --.

Column 35
Line 3, "A1107" should read --A107--.

Line 9, "and complements" should read --and the complements--.

Line 13, "A55A56A59A92A94to" should read --A55, A56, A59, A92, A94 to--.

Line 13, "A108A111 to A113A115" should read --A108, A111 to A113, A115--.

Column 36
Line 13, "^{32}p" should read --^{32}P--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 38
Line 45, "VLSIPSυ" should read --VLSIPS™--.

Column 50
Line 16, "A123 complements" should read --A123 to A125, and the complements"--.

Line 20, "A114, A122" should read --A114, and A122--.

Line 25, "to 113, A115, A117, and" should read --to A113, A115 to A117, and--.

Line 64, "A123 complements" should read --A123 to A125, and the complements--.

Column 51
Line 1, "A109, A122," should read --A109, A110, A114, and A122--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page

28 JAN 2005

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 52
Line 10, "are Q selected" should read --are selected--.

Column 56
Line 48, "the w biallelic" should read --the biallelic--.

Column 58
Line 25, "A109, A122" should read --A109, A110, A114, and A122--.

Line 29, "A94 to A01," should read --A94 to A101,--.

Column 62
Line 9, "Eli" should read --E11--.

Column 66
Line 25, "about IS" should read --about 15--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies

@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 67
Line 27, "WO 98120165" should read --WO 98/20165--.

Column 73
Lines 22-23, "with overlapping" should read --with non-overlapping--.

Column 87
Line 2, "(a_i,a_j, a_i,b_j: b_i,a_j)" should read --(a_i,a_j; a_i,b_j; b_i,a_j)--.

Line 7, "θ4=__=frequency" should read --θ4= - - = frequency--.

Line 9, "θ3=–+=frequency" should read --θ3= - + = frequency--.

Line 11, "θ2=+–=frequency" should read --θ2= + - = frequency--.

Line 14, "(M_i; M_j)" should read --(M_i, M_j)--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192
No. of add'l. copies
@ 30¢ per page

28 JAN 2005

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 87
Line 21, "D_{aiaj} ($2n_1$+" should read --D_{aiaj}= ($2n_1$+--.

Column 92
Line 29, "AS5" should read --A55--.

Column 95
Lines 41-42, "preferred 1 5 embodiment" should read --preferred embodiment--.

Column 101
Line 49, "be I,X;1 expressed" should read --be expressed--.

Column 109
Line 22, "one-genes" should read --*onc*-genes--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies

@ 30¢ per page

FORM PTO-1050 (REV. 3-75) UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 117
Line 15, "(ISPR)" should read --(SPR)--.

Line 17, "carboxymethyi" should read --carboxymethyl--.

Column 118
Line 16, "(MaTa gal4" should read --(MATa gal4--.

Line 18, "lacZmet')," should read --lacZmet-)--.

Line 52, "GALA" should read --GAL4--.

Column 125
Line 6, "more CF particularly" should read --more particularly--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 127
Line 57, "monoleatc" should read --monoleate--.

Column 132
Line 42, "CAROM" should read --CD-ROM--.

Column 134
Line 47, "or a By polypeptide" should read --or a polypeptide--.

Column 136
Line 47, "TTATAA Box" should read --TAATAA Box--.

Line 48, "TATAA" should read --TAATAA--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies

@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 140
Line 10, "[BLANK]" should read --PCR assays were performed using the following protocol:--

Column 158
Line 1 of Table 6, "analysis of the" should read --analysis for the--.

Column 159
Line 2, "analysis of the" should read --analysis for the--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 159
Line 6 of Table 6,
"

	A2	A30	A41	A55	A57	A75

"

should read

--

	A2	A30	A41	A55	A57	A75
frequency %	60/67 (A)	64/66 (T)	73/71 (T)	64/68 (C)	65/69 (G)	94/95 (G)
abs diff freq. all.	-7,4	-2,0	2,7	-3,7	-3,9	-1
pvalue	7,7e-03	4,3e-01	2,9e-01	1,6e-01	1,4e-01	3,4e-01
Cases/controls ↓						

--

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192
No. of add'l. copies
@ 30¢ per page

FORM PTO-1050 (REV. 3-75) UNITED STATES PATENT AND TRADEMARK OFFICE

28 JAN 2005

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 164
Line 14, "and b dose" should read --and dose--.

Line 34, "they arc" should read --they are--.

Column 167
Line 25, "Baltimore, 1991 Ouchterlony, O." should read --Baltimore, 1991 Ouchterlony, O.--

Column 168
Line 22, "119-1123" should read --1119-1123--.

Line 26, "Stembert" should read --Sternberg--.

Line 26, "5:397404" should read --5:397-404--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,759,192

DATED : July 6, 2004

INVENTORS : Marta Blumenfeld, Lydie Bougueleret, Ilya Chumakov

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 168
Line 27, "Strinmatter" should read --Strittmatter--.

Column 569
Line 28, "and b) performed" should read --and b) are performed--.

Column 570
Line 25, "A75 and A30;" should read --A75;--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
2421 N.W. 41st Street, Suite A-1
Gainesville, FL 32606

PATENT NO. 6,759,192

No. of add'l. copies
@ 30¢ per page